

Who are we?

Efforts to catalogue and understand the human microbiome are opening up a whole new research frontier. But the earlier Human Genome Project should provide a cautionary lesson about overselling.

“Who am I?” is a question that is often asked and seldom answered. But as several articles in this issue suggest, the question itself may need to be reframed: biologists are discovering that it is frequently more informative to ask, “Who are we?”

The ‘we’ refers to the wild profusion of bacteria, fungi and viruses that colonize the human body. These unseen passengers number in the trillions. According to one common estimate, the human gut contains at least a kilogram of bacteria alone. They contribute so much to human biology that it is difficult to say where the body ends and the microbes begin — which is why several massive projects have now started up to characterize the human microbiota in its entirety (see page 578).

Microbiologists are understandably excited by this opportunity. So, too, are the food and pharmaceutical industries. When it comes to profitable applications to human health, the microbiome could well offer distinct advantages over the more famous genome. Human genes are notoriously difficult and risky to tamper with. But, in theory at least, the microbiome should be relatively easy to change by the selective addition or removal of bacterial species, or by altering their genetic components. This idea has some basis. Antibiotics and ‘probiotic’ foods have already been shown to calm inflammatory bowel diseases in some instances. In this issue, for example, researchers show how intestinal inflammation can be reduced by a single molecule produced by a gut bacterium (see pages 602 and 620). And there is increasing acceptance that certain foods, or the bacteria contained in them, can alter gut microbiota in ways that are beneficial to health in general.

The new appreciation of the microbiome comes just as some observers have started to question whether the human genome can deliver on its once-hyped promises to tackle disease. To take just one

example, anyone so inclined can now pay genetic-testing companies for a preliminary rundown of the genetic variations associated with his or her risk of developing cancer, obesity and other conditions. But the risks identified are often so low or unclear that people are questioning whether the information will actually prompt the changes in health behaviour, such as losing weight, that could make them valuable (see page 570).

For all the excitement, however, researchers involved in the human microbiome efforts can learn a valuable lesson from the genome experience. Simply put: be circumspect. Don’t oversell the human microbiome until its medical promise has been established. Remember that the understanding of these microbial communities is still fragmentary, at best — and that it is far from established that the microbiota can be radically altered without upsetting the balance and causing harm, or that any alterations will last more than a few months. Indeed, attempts to understand the dynamics of gut colonization are still in their infancy (see page 581).

“Don’t oversell the human microbiome until its medical promise has been established.”

In the meantime, microbiologists should celebrate their quest to map, catalogue and understand the human microbiome for the inspiring saga it is. Certainly there is food for thought in the fact that everyone has inside them exotic environments that support communities as diverse as any rainforest. There is a unique ecological perspective on food itself, and the effects that different foodstuffs, such as processed versus unprocessed ones, have on these environments.

There is a compelling new take on humankind’s place in the world — a realization that “Who am I?” cannot be fully answered until it is fully understood who ‘we’ are. ■

Animal tests inescapable

The ambitious scope of Europe’s chemicals legislation demands some innovative toxicology.

The European Union’s (EU) Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) directive, which took effect a year ago, on 1 June 2007, is widely regarded as the strictest chemical safety law in the world. Unlike the 1981 EU legislation it replaces, or the US Toxic Substances Control Act, REACH applies to all existing chemicals, not just new ones.

If nothing else, that requirement is likely to cause six months of immense stress at the organization created to administer REACH, the European Chemicals Agency in Helsinki. The agency will open its doors to registration on 1 June this week. And with every chemical

manufacturer or importer in Europe facing a 1 December deadline to do the initial registration on every compound it had put on the market before 1981, agency staffers are braced for an estimated 180,000 applications.

More importantly, however, the requirement to test all those old chemicals has made REACH a case study in difficult choices. In following one policy — ensuring the safety of its citizens — the European Union may have severely strained another: lessening the use of animals in regulatory toxicology, and in research in general. Working through that backlog without sacrificing a lot of animals will be impossible.

Happily, the actual number of sacrifices is likely to be less dramatic than it might have been. Partly this is because safety data are already thought to exist on many of the older chemicals (although no one has comprehensive statistics). And partly it is because of REACH itself. In 2002, while the directive was still being drafted, the European Union expanded and accelerated work at the European Centre for

the Validation of Alternative Methods: a small laboratory in Ispra, Italy, devoted to finding kinder and more-refined toxicity tests that would require fewer, or no, animals. That work has begun to pay off. For example, the notorious LD₅₀ test is required, which identifies the dose that kills half of the tested animals, now involves around ten rats rather than 45. And instead of looking for a dose that causes outright (and painful) skin allergy in guinea pigs, toxicologists can now monitor mice for the painless swelling of lymph nodes, the first sign of allergy.

These are the only two animal tests required for chemicals produced in relatively small quantities (1–10 tonnes a year). Those produced in smaller quantities require no animal testing at all. Not so happily, however, the refinements will have little effect when it comes to chemicals produced in larger volumes. REACH will require many more tests on such compounds, including, in some cases, their effects on reproduction and fertility, which use 3,000 animals per chemical.

Linking up development

An initiative to link scientists in the poorest nations with colleagues around the world deserves support.

In September 2000, the United Nations pledged to achieve eight Millennium Development Goals by 2015 — among them, halving extreme poverty, providing primary education to all and stopping the spread of HIV/AIDS. The goals have spurred unprecedented efforts by countries and development agencies alike. But some of those efforts, such as the controversial One Laptop Per Child Project, which is widely seen as a developed-world solution to a developing-world issue, should also serve as a warning of the illusion that the problems of development can be solved with technological quick fixes.

That circumspection should be borne in mind by Scientists Without Borders — an initiative launched with great fanfare on 12 May by the New York Academy of Sciences in partnership with the UN Millennium Project. The initiative aims to address health and other issues in developing countries by using the Internet to bring together scientists from different countries and fields.

The academy argues that the initiative will “create a broad array of synergistic linkages among the many bold but heretofore unconnected efforts to generate science-driven, sustainable development in the poorest of the poor communities”. As a result, it predicts, universities in rich and poor nations will volunteer to work together to train local staff, and people and institutions will learn from one another and collaborate on development issues in a more interdisciplinary and integrated fashion.

The caveat, of course, is that better communication between scientists alone cannot magically transform Africa, or any other part of the developing world. And although the virtual Scientists Without Borders has gained the support of some impressive partners — including France's Institut Pasteur and the American Society for Cell Biology in Bethesda, Maryland — it has only one full-time staff member, and has raised US\$1 million. As such, it has little in common — aside from

Beyond REACH, moreover, the 2003 amendment to the EU Cosmetics Directive will soon add even more time pressure: an end to acute toxicity testing of cosmetics (including soaps and shampoos) in animals by 2009, and the longer and more involved repeat-dose testing by 2013. Scientists consider the latter deadline to be hopelessly unachievable. There is simply no way in the foreseeable future that they will be able to model and predict all the ways that a foreign molecule might interact over time with the hundreds of types of tissues in the human body.

Yet that is all the more reason to put some serious effort into tackling the problem. Revolutionizing regulatory toxicology won't be cheap and it won't be quick, although systems biology — the integration of genomics, proteomics and all the other ‘-omics’ — could prove to be a particularly powerful ally. But it should be a priority. And, as shown by the cosmetics directive's unrealistic deadline (fortunately renegotiated in 2011), scientists have to be brought into policy decisions at the beginning, and listened to carefully. ■

its name — with the operational agency Doctors without Borders (Médecins Sans Frontières), which has a €569 million (US\$ 896 million) annual budget, and thousands of staff around the world.

That said, however, we hope that the academy is right, and that Scientists Without Borders will rapidly become the Facebook of development. The initiative certainly deserves scientists' full support for its efforts to boost the capacity of researchers in poorer countries. Consider its first initiative: to create an online database in which scientists and institutions can register their profiles, and so connect with others who share similar interests. Yes, many databases and social networks already exist for scientists. But Scientists Without Borders is unique in its commitment to raise the visibility of research in developing countries, and to compile an inventory of who is doing what — specifically limited to those who have expressed an interest in cooperating. This in itself is a worthwhile exercise.

The database will also tap user profiles to match research needs with available resources, say, or to announce the availability of volunteer work and research collaborations — although it is still far from clear how this will work in practice.

Scientists Without Borders faces a wider challenge given its coverage of all disciplines and sectors, as it makes it difficult for experts to pick out the information that is relevant to them. Online communities of practice tend to flourish around bottom-up approaches that generate centres of common interest that keep members active and coming back for more. But it is encouraging that the organizers seem keen for users to suggest changes to the Scientists Without Borders platform to meet their needs. Experience elsewhere shows that demand for social networks is critically dependent on them satisfying real user exigencies.

Most importantly, the initiative has generated hope. As one new user told *Nature*, it has for the first time provided him with a space to “put my skills on the radar of those directly involved in projects in the developing world; I hope I can contribute in some way.” We would encourage scientists to sign up and participate, and nurture this fledgling initiative. ■

“Scientists should nurture this fledgling initiative.”

RESEARCH HIGHLIGHTS

**Pretty good***Ecology* **89**, 1201–1207 (2008)

There is strength in diversity, say Anders Forsman and Viktor Åberg of the University of Kalmar in Sweden. The pair looked at 323 species of Australian

reptile — 275 lizards and 48 snakes — and found that species that exhibit variable skin patterns have ranges almost three times larger, on average, than do reptile species with less variable colouration. They also live in more habitat types.



The researchers propose that multiple colour patterns reflect combinations of traits that, by evolving together, allow a single species to exploit a range of environments. Fifty of the species studied are classified as threatened, and of those

only one (the death adder, *Acanthophis antarcticus*) showed variable coloration (pictured). So this finding may provide conservationists trying to gauge the status of little-surveyed reptiles with a useful rule-of-thumb.

S. D. MILLER/R. VALENTIC/NATUREPL.COM

CELL BIOLOGY**Bean counting***Cell* **133**, 627–639 (2008)

A newly identified protein complex allows cells to vary the rate at which they make ribosomes — the factories that translate RNA into protein — in response to how much energy is available. Ribosome production is the most energetically expensive thing that eukaryotic cells do, so the complex has an important role in cell survival.

A team led by Junn Yanagisawa of the University of Tsukuba in Japan isolated a protein complex called NoSC that controls transcription of ribosomal RNA, the main component of ribosomes, in response to the ratio of NAD⁺ to NADH molecules, a signal of the available energy in a cell. NoSC contains a previously uncharacterized protein, named nucleomethylin, that binds one of the histone proteins with which DNA is packaged. It also comprises SIRT1, a protein made in response

to caloric restriction and SUV39H1, which can alter gene expression.

THEORETICAL PHYSICS**Better out than in***Phys. Rev. Lett.* **100**, 21302 (2008)

Surviving a black hole is not easy, but calculations by Abhay Ashtekar of Pennsylvania State University, University Park, and his collaborators suggest that under certain conditions quantum information might make it out.

This team of physicists imagines space-time as having a fundamentally quantum structure. Considered in this way, a black hole's point of infinite mass and gravitational pull, known as its 'singularity', disappears, and quantum fluctuations can travel right through the black hole's core.

This result is important because it fulfills the stipulation of quantum mechanics that information is always conserved. The authors

hope their work might one day help to integrate quantum mechanics with theories of gravity.

GENETICS**Tiger-mice***PLoS One* **3**, e2240 (2008)

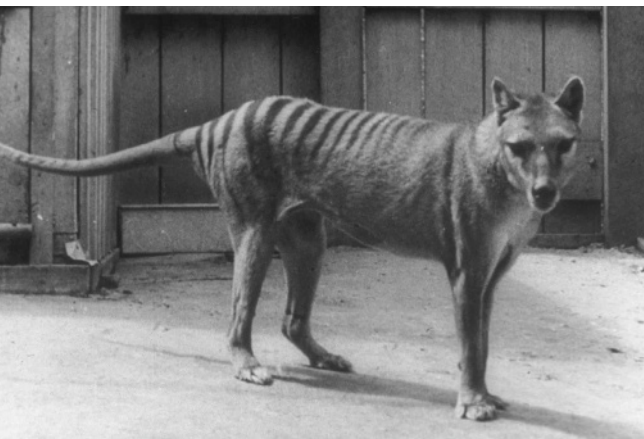
DNA from the Tasmanian tiger (*Thylacinus cynocephalus*, a marsupial, pictured left) has been put to work inside mice, marking the first example of successful genetic transfer from an extinct species to a living one.

Richard Behringer of the University of Texas M. D. Anderson Cancer Center in Houston and his colleagues extracted DNA from a century-old Tasmanian tiger skin and three preserved pouch young held in the Museum Victoria in Melbourne. The researchers then took a regulatory part of the Tasmanian tiger's *Col2a1* gene, which is involved in the development of bone and cartilage, and introduced this into the genomes of several mouse embryos. These modified mice grew up normally, with the inserted section of tiger-*Col2a1* functioning in their developing skeletons.

NANOTECHNOLOGY**Nanozapped***Nano Lett.* doi:10.1021/nl080661a (2008)

Tiny, levitating semiconductor particles can produce laser light. This happened when a group of researchers led by Lijun Wang at the University of Erlangen in Germany encased quantum dots with cadmium-selenide cores and zinc-sulphide shells in microdroplets of water and glycerine, and electrically charged the droplets to keep them floating in air. When energized with light from a laser, the encased quantum dots generated their own laser light.

A microdroplet forms a super-smooth spherical capsule around the quantum dots, causing the photons they give off to oscillate at specific wavelengths, a requirement for lasing. Surprisingly, the density of quantum dots within a single microdroplet needed to produce laser light was very low, the researchers say — as was the power needed to drive the process.



D. WATTS/ALAMY

PHYSICAL CHEMISTRY

Growing up bigger

Angew. Chem. Int. Ed. **47**, 4208–4210 (2008)

Which is smaller: hydrogen or deuterium? The standard answer is that deuterium (^2H) takes up less space than ^1H because its greater mass gives it a smaller vibration amplitude in the lowest-energy quantum state. But the veracity of this argument is temperature sensitive, say Jack Dunitz of ETH-Zurich in Switzerland and Richard Ibberson of the Rutherford Appleton Laboratory in Didcot, UK.

By taking careful measurements of the crystal structures of benzene and fully deuterated benzene at temperatures between 5 kelvin and 280 kelvin, they show that the volume of a molecule of the latter exceeds that of the former above about 170 kelvin. At such temperatures, the vibrations of the carbon–deuterium bonds include more of the higher-energy quantum states than those of $\text{C}-^1\text{H}$ bonds, which means that the deuterium atoms effectively occupy more space.

GENETICS

Genes with bottle

Nature Genet. doi:10.1038/ng.151 (2008)

Researchers have pinpointed a pair of gene variants that seem to protect their carriers against head and neck cancers linked to alcohol consumption. The variants are of genes that encode alcohol dehydrogenase (ADH) enzymes, which catalyse the breakdown of alcohols into aldehydes.

Paul Brennan at the International Agency for Research on Cancer in Lyon, France, and his colleagues analysed six ADH variants in 3,876 patients with head and neck cancers and 5,278 healthy people. A variant of each of the two genes *ADH1B* and *ADH7* seemed to lower the risk of developing these cancers in alcohol

drinkers, most dramatically in heavy drinkers. Carriers of the *ADH1B* variant metabolize alcohol up to 100 times faster than non-carriers.

GEOPHYSICS

The heat is on

Earth Planet. Sci. Lett. doi:10.1016/j.epsl.2008.03.031 (2008)

A planet's interior affects its climate through volcanoes spewing out greenhouse gases. Conversely, the climate can also affect the interior, according to calculations by Adrian Lenardic, of Rice University in Houston, Texas, and his co-workers.

They worked out that temperature increases at a terrestrial planet's surface could penetrate deep into the planet, rendering its mantle less viscous and eventually shutting down the movement of tectonic plates.

For a planet such as Earth, a sustained rise of 100 kelvin over a 10-million to 100-million-year timescale could be enough to destabilize plate tectonics. The authors suggest that the carbon dioxide blanket in Venus's atmosphere (artist's impression, pictured) might help to explain why it appears to have a single, static plate.

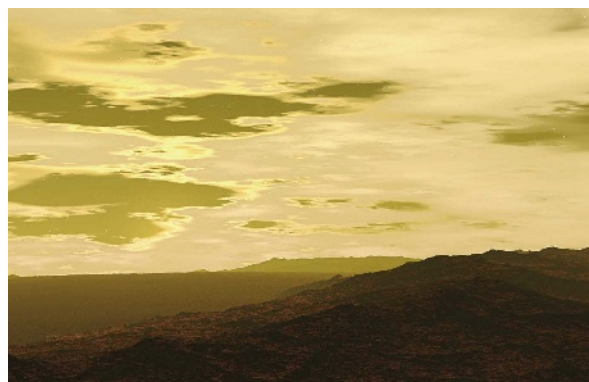
GENETICS

The sweet life

Hum. Mol. Genet. doi:10.1093/hmg/ddn137 (2008)

Different versions of the gene that encodes insulin-degrading enzyme (IDE) are associated with how long men — but not women — live, researchers have found.

Insulin metabolism had previously been linked to the lifespans of organisms commonly used in laboratory research. Jonathan Prince of



C. DANKIN/SPL

the Karolinska Institute in Stockholm, Sweden and his colleagues now report that human males with one copy of a certain version of IDE and one copy of a different version make higher than usual amounts of IDE RNA and produce more insulin when they fast. These men also tended to die younger than those carrying two copies of the IDE variant.

MICROBIOLOGY

Inner lives

Science doi:10.1126/science.1155725 (2008)

How particular to humans are the microbes that eek out a living in our intestines? To find out, Jeffrey Gordon of Washington University School of Medicine in St Louis, Missouri, and his colleagues compared particular genetic sequences from gut microbes found in the faeces of 106 mammals, some wild and others from American zoos.

Among the 60 species represented in this sample, carnivores tend to have a less diverse internal flora than omnivores, and omnivores less than herbivores. Modern humans support similar gut microbes to those of other omnivores — a surprise, maybe, given the importance of agriculture and cookery in human ecological history.

JOURNAL CLUB

Nathan Wolfe
University of California, Los Angeles

An epidemiologist points to a fifth sort of human malaria.

Malaria has plagued humans since the dawn of written history, and probably since long before that. These days, biologists understand tiny mechanistic details of the workings of one human malarial parasite, *Plasmodium falciparum*, but know surprisingly little about the others. As someone who

studies how pandemics are born and die — and how they might one day be prevented — these holes in our knowledge seem striking to me.

Aside from *P. falciparum* — the cause of 'malignant' malaria — parasitologists acknowledge three other human malaria parasites, *P. vivax*, *P. ovale* and *P. malariae*, each of which probably jumped from another primate host to humans independently. With so many malaria parasites plaguing other vertebrate species, however, and only basic diagnostic instruments available in most parts of the world, science could

be missing new types of human malaria that have the potential to seed pandemics.

In a recent paper, Janet Cox-Singh and her colleagues build on their earlier finding that humans can harbour a fifth malaria parasite, *P. knowlesi*, which was once thought to infect only Asian monkeys. The researchers detected *P. knowlesi* DNA in about one third of 1,014 malaria patients in Malaysia, showing that this parasite is common, deadly and almost always misidentified as *P. malariae* (J. Cox-Singh *et al. Clin. Infect. Dis.* **46**, 165–171; 2008).

That an unknown animal pathogen can cause widespread human disease is reminiscent of some of the biggest scourges of the twentieth century: HIV and pandemic influenza. Reductionist, molecular approaches to tackling important plagues may be *en vogue* and a near necessity for grant funding, but I bet that an old-fashioned natural historian studying how infectious agents jump host species will be first to signal the coming of the next plague.

Discuss this paper at <http://blogs.nature.com/nature/journalclub>

NEWS



Venerable institute gets a refit

The oldest independent research body in the world, Britain's Royal Institution, has been committed since 1799 to "diffusing science for the common purposes of life". But some of the common purposes of life have changed in the past two centuries.

Visitors to the Royal Institution in London's Albemarle Street — which was once so besieged by carriages attending the popular scientific lectures that it was made the city's first one-way street — now require a more modern, broader variety of engagement, the Royal Institution's director believes. The palatial building, which has hosted exciting events such as the discovery of sodium and potassium by Humphrey Davy and electromagnetic induction by Michael Faraday, became run down and dilapidated over time.

So, time for a change. This week, after a £22-million (US\$44-million) refurbishment, the building — part public venue, part research laboratory — reopened to acclaim. A two-year restoration project to preserve the historic fabric of the building has seen major changes to public access and facilities. A café, bar and a hologram of Faraday now greet visitors to an institution that once struggled to offer the public so much as a cup of tea.

The aim, according to director Susan Greenfield, "is to position the Royal Institution at the centre of twenty-first-

century society, where science will take its proper place at the heart of all our lives". Greenfield will retain her private apartment at the top of the building, a privilege of institution directors.

The Royal Institution is perhaps best known for its Christmas lectures, annual science talks for children that were started by Faraday in 1826 and reach a worldwide audience. But the institution is in a rare position of combining public engagement in science with a tradition of world-class research that has seen the discovery of ten chemical elements and bagged 14 Nobel prizes. A new team will soon settle into the institution's Davy-Faraday Research Laboratory (DFRL), where scientists pioneered X-ray crystallography and the vacuum flask.

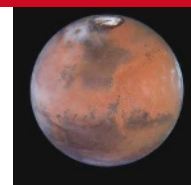
Greenfield, a neuroscientist who was appointed director in 1998, four years after becoming the first woman to deliver the Christmas lectures, has brought a new vision to the Royal Institution and increased the 'public engagement' operations. The constituency attending its events is now much wider than the academic and old-boys' networks of old. A lecture hall and cloakroom no longer suffice. The modern audience demands food, drink, entertainment and education from the same venue.

Architects at Terry Farrell and Partners have redesigned the building to meet these new needs. Walls have been knocked through to make way for a glass atrium and restaurant space. The museum sections of the basement have been refreshed to celebrate the institution's

history and heritage. The Faraday lecture theatre, home to the Christmas lectures, is completely refurbished (see left). The venue can even be hired for weddings.

But although public access is improved, the number of scientists in the DFRL has been reduced from 60 to 15. When the site closed for renovation, the laboratory moved temporarily to University College London (UCL). In the intervening two years, the group, headed by Richard Catlow, became "increasingly dependent on UCL's facilities" and has chosen to remain at UCL, a decision that all parties describe as amicable.





PHOENIX LANDS NEAR MARS NORTH POLE

Get blog updates on the search for signs of habitability.

www.nature.com/news

But a former member of the laboratory, who wishes to remain anonymous, feels that the research group was increasingly squeezed out in favour of public engagement even before the refurbishments. He questions Greenfield's vision, and whether it was worth the institution selling its property assets to finance the changes. "Greenfield was right, the Royal Institution did need to change. But whether she was right about what it should change into is an open question that will be answered in the coming months," he says. "The redevelopment has cost the entire £15-million property portfolio and the stakes are extremely high."

Last month, Quentin Pankhurst, deputy director of the London Centre for Nanotechnology, was appointed Catlow's successor. As director of the DFRL he will oversee a team of 15 permanent staff, including three principal investigators as well as several PhD students and postdocs. "I want to create a highly interdisciplinary, close-knit unit here," says Pankhurst, who will retain an affiliation with UCL, which is jointly financing the lab with the Royal Institution. Pankhurst's group will explore health-care biomagnetics, using nanopar-

"We want to show that scientists are human too."

ticles attached to antibodies to locate and treat cancers — techniques that hark

back to Faraday's electromagnetic-induction experiments in the early days.

The group will take up 750 square metres of lab space, roughly a third of that available. The rest will go to small start-up companies not affiliated with the DFRL.

Part of Pankhurst's lab will be on public view. "We want to show that scientists are human too," he says. "The open lab won't be populated minute-by-minute, but visitors will see Mössbauer and X-ray techniques as and when the scientists use them."

Pankhurst dismisses concerns that research will play second fiddle to outreach programmes. The Royal Institution's mission statement from 1799 compels it "to facilitate the general introduction of useful inventions and improvements; and to teach... the application of science for the common purposes of life". According to Pankhurst: "These two aspects, which we might now call technology transfer and public engagement, have been and still are central tenets of the Royal Institution." ■

Matt Brown

See Books and Arts, page 595.

Climate anomaly is an artefact

The humble bucket turns out to be at the bottom of a perplexing anomaly in the climate records for the twentieth century.

The time series of land and ocean temperature measurements, begun in 1860, shows a strange cooling of about 0.3 °C in the global mean temperature in 1945, relative to the 1961–90 average. The sharpness of the drop stands out even more if the signatures of internal climate variability, such as those associated with El Niño events, are filtered from the record.

This cooling at the end of the Second World War is one of several temperature drops in the record. But unlike others, such as the 1991 cooling caused by the eruption of Mount Pinatubo in the Philippines, it is limited to ocean temperatures and is not associated with any known climatic or geological phenomenon. The nuclear explosions in 1945 over Hiroshima and Nagasaki were ruled out as a possible cause because they are thought to have had no impact on global temperature. Other theories proposed as explanations for the cooling include a massive 1940s El Niño event that had somehow slipped attention, or that it was the result of sulphate aerosols from burning dirty coal. But neither of these was convincing.

A US–British team of climate scientists has now found a surprisingly simple explanation for the long-standing conundrum (page 646). It turns out that the mysterious drop is due to differences in the way that British and US ships' crews measured the sea surface temperature (SST) in the 1940s.

Only a few SST measurements were made during wartime, and almost exclusively by US ships. Then, in the summer of 1945, British ships resumed measurements. But whereas US crews had measured the temperature of the intake water used for cooling the ships' engines, British crews collected water in buckets from the sea for their measurements. When these uninsulated buckets were hauled from the ocean, the temperature probe would get a little colder as a result of the cooling effect of evaporation. US measurements, on the other hand, yielded slightly higher temperatures due to the warm engine-room environment.

The standard logbook entries made at the time contain no information about how the measurements were taken, so the cause was overlooked, says David Thompson, first author on the paper and an atmospheric scientist at the State University of Colorado in Boulder. As a result, the bias — which, although small, was large enough to produce the sharp drop in global mean temperature — was never adjusted for.

"The time series is one of the great climate records we have," Thompson says. "During a sabbatical in Britain, I revisited work that I had

started a long time ago, and it suddenly occurred to me that the mid-1940s cooling might not necessarily have physical causes."

Thompson discovered the explanation after questioning maritime experts from different countries about the history of shipping, and searching the scientific literature and international databases for scattered bits of relevant information.

"We always thought the observed cooling was real," says Phil Jones, a climate researcher at the University of East Anglia, Norwich, UK, who carried out the study with Thompson. "We did know that there were fewer measurements during the war than before and thereafter, but we simply made wrong assumptions on how and by whom the measurements were taken," he says. "It is pretty clear now that the bias is instrumental."

"It is welcome news for climate modellers. The post-war temperature anomaly has been grossly outside the range of all computer-based climate reconstructions considered by the Intergovernmental Panel on Climate Change (IPCC), and it was prominently featured in the group's 2007 summary for policy-makers."

"The unusual up and down in SSTs in the 1940s stood out like a sore thumb in the past," says Susan Solomon, a senior scientist at the US National Oceanic and Atmospheric Administration (NOAA) in Boulder, Colorado, and co-chair of the IPCC working group on the physical basis of climate change. "We couldn't explain it, so we showed all the fingers, sore thumb and all," she says.

Climate researchers can now start setting the twentieth-century temperature record straight. The abrupt drop in 1945 will then probably disappear, but what the corrected time series will look like is not yet clear.

And further corrections may come. For example, the gradual shift since the 1970s from (warm-biased) ship-based measurements to (cold-biased) drifting buoys has probably led to a slight underestimate of SST warming, says Richard Reynolds of NOAA's National Climatic Data Center in Asheville, North Carolina. "More data problems will undoubtedly come along, if likely of smaller magnitude," he says.

Climate researchers know from past experience that satellite sensors, radiosondes and ocean profilers are prone to bias. Land measurements are considered much more robust. Climate scientists should think about data quality more often, says Jones, so that there is no opportunity for incorrect data to sow seeds of doubt in people's minds about the reality of climate change. ■

Quirin Schiermeier

SPECIAL REPORT

Genetic testing for everyone

Private companies are starting to test customers' DNA for gene variants linked to an increased risk of conditions such as obesity or Alzheimer's disease. **Helen Pearson** looks at whether knowledge really is power when it comes to disease avoidance.

Terry Drotos was adopted at birth. She knows nothing of her biological family's medical history and so cannot use it to gauge her or her daughter's future health. When she was offered the chance to have her genetic information analysed, she jumped at it.

Drotos was one of the first to try the genetic testing service offered by Navigenics, based in Redwood, California. After the company analysed her DNA for gene variations linked to 19 conditions, she was relieved to hear that her susceptibility to Alzheimer's disease is no higher than average. But she was surprised to learn that her risks of obesity and type 2 diabetes are 2–3% higher than average for her age group. Drotos is slim and exercises regularly, so there was little more she could do to lower those risks. "It really does show that if I wasn't already taking care of myself, I probably would be overweight or diabetic," she says, although geneticists would probably disagree with her assessment.

"It's an intriguing idea that you can peel back your genome and reveal your future."

Direct-to-consumer genetic testing is a rapidly growing market — the past year has seen the launch of companies, such as Navigenics and 23andMe in California and DeCODEme in Iceland, that offer DNA screening for a range of common genetic variants linked to disease. The testing outfits have created a buzz in the business and research communities as well as in the wider public: Google has invested in two of them and Navigenics briefly opened a store in New York's hip SoHo district.

The idea is that test-takers will be alerted to risks and so take preventive action where possible. But psychosocial scientists who study how people respond to risk information say there is scant evidence that people are affected deeply by genetic test results, or that such tests spur much change in behaviour.

"In public these genetic-testing companies have made a big splash, and it's an intriguing idea that you can peel back your genome and



Will genetic testing add anything when it's already known that obesity is a health risk?

reveal something about your future," says Robert Green, who studies Alzheimer's disease at Boston University in Massachusetts. "But if that idea disappoints," he adds, "some of the lustre will fade."

The service provided by these companies

Breaking the news

A common perception is that hearing the results of a genetic test confirming an increased risk of, for example, breast cancer, causes shock, distress and lasting psychological scars. So genetic information is treated as medical dynamite. This is knowledge that must be communicated gently, the conventional wisdom goes, preferably by a trained genetic counsellor.

No one denies that genetic test results can be life-altering for some individuals. But research by Theresa Marteau, a health psychologist at King's College London, and others has shown that most people are remarkably

resilient in the face of traumatic genetic test results. They typically report feeling anxious or depressed around the time of testing, but these effects dwindle within a few months. Studies by Aad Tibben, a psychologist and psychotherapist at Leiden University Medical Centre in the Netherlands, and his colleagues showed that people who took predictive tests for Huntington's disease mostly recovered from the shock. Many actually felt more in control after testing because they could make arrangements for care, or even for euthanasia.

With so much uncertainty about how people deal with genetic risk,

is genetic counselling necessary or helpful for people undergoing the less definitive tests for an increased propensity for heart conditions or diabetes? "I'm convinced it's necessary," says Tibben. But he and others in the field acknowledge that there is little in the way of controlled trials to support their belief.

"We can't say empirically that people are better off after seeing a genetic counsellor," says Barbara Biesecker of the US National Human Genome Research Institute in Bethesda, Maryland. Even if counselling does seem beneficial, as some studies suggest, it is not known

why. Did the counsellor help the patient understand complicated risks, or just provide some face-to-face contact and empathy in a confusing medical world?

Studies are under way to answer some of these questions. California-based Navigenics is co-sponsoring a large, five-group trial with researchers at the Mayo Clinic in Rochester, Minnesota. This will examine how participants react psychologically and behaviourally to medical risk information based on different sources, including family history and genetic testing, and presented to them with or without counselling. **H.P.**



differs from most existing genetic testing, which has been a niche medical service limited to those with a family history of breast or colon cancers or other hereditary diseases. A positive genetic test for these conditions means that the person is at very high risk or even, in the case of the neurodegenerative Huntington's disease, that the condition is inevitable.

Changing behaviour

Very little is known about how people will respond to the new genetic tests, which reveal a small increased susceptibility to complex conditions involving many genetic and environmental factors. For example, Drotos was told that her lifetime risk of obesity was 34% compared with an average of 32% for someone of her sex. In this sense the tests are more akin to finding out about high blood pressure or excess weight, factors that are linked to heart disease and diabetes but that do not say for sure whether any individual will get them. "It would be silly to think these small risks will have a major psychological impact," says Barbara Biesecker, a genetic counsellor at the US National Human Genome Research Institute in Bethesda, Maryland.

One of the very few clinical trials to examine how recipients handle this type of low-risk information is being led by Green.

Half of the several hundred participants were assigned to receive genetic testing for a variant of the apolipoprotein E (ApoE) gene. People with two copies of this variant have an increased risk of Alzheimer's of up to 15 times the average for the general population. Green says that "people do very well with this information", with any mild initial distress (see 'Breaking the news') being resolved within about six months after being told. His team has also found that those with the risk variant were more likely to

make changes to their diet or medication even though there is no proven way of preventing Alzheimer's.

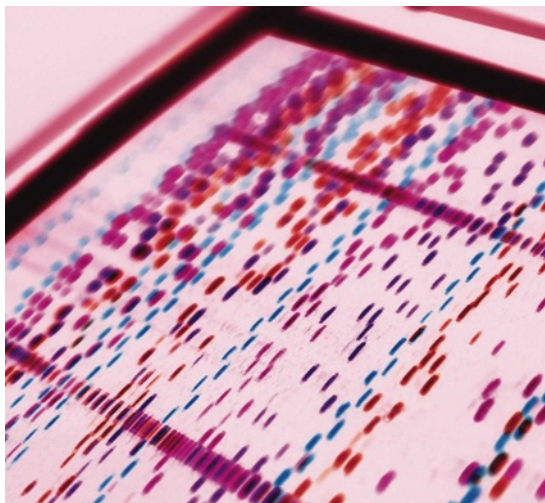
Many researchers question whether genetic information will be any better than existing health warnings or family medical history at prompting people to change their behaviour. Most people with a generous waist measurement know they're at risk of obesity-related conditions and what they ought to do about it — genetic information may not help them.

"The question is whether genetic information is a big enough motivator," says Susanne Haga, who studies health policy and genomics at Duke University in Durham, North Carolina. "It's naive to think people are going to have a complete turnaround in their lifestyle." Indeed, one concern is that testing could backfire: people who find out they are at low risk for a condition could use this as an excuse not to adopt healthy behaviour. "It could go the opposite way as well," Haga says.

Elissa Levin, director of genetic counselling for Navigenics, says that genetic test information may help people to refine their health advice in combination with a family history or other risk factors — perhaps spurring them to have earlier screening for colon cancer. "There is a lot to say about the utility of this information apart from 'exercise and eat well,'" she says. "We'd be pretty surprised if there was no benefit." Although Drotos was already doing all she could to ward off obesity, she says her test results prompted her daughter to shed some weight.

With so little hard evidence that genetic testing is beneficial, many experts question whether there is much value at all for people who buy the new tests. "The most likely outcome would be disillusionment that they paid a lot of money and got very little," Biesecker says. But more is at stake for healthcare policy-makers, who have to decide whether paying for genetic testing or counselling, now or in the future, will give a return in improved public health. "I don't think healthcare has faced something quite so challenging as this," says David Veenestra, who studies health policy at the University of Washington in Seattle.

Veenestra believes that the value of genetic information may go up as the cost of testing and sequencing drops, and if, as he predicts, trials begin to show that genetic testing can cause small changes in behaviour, refinements in drug dosing or other benefits. Even if only 5% of people change their lifestyle for the better after testing, that could be worth the cost over a population. "If you start to add those up, your initial investment would start to look worthwhile," he says. ■



Genetic tests can only predict one risk factor for common conditions such as heart disease and diabetes.

F. DEAN/REX FEATURES

A. BROOKES/CORBIS

News

France's research agency splits up

The CNRS is being carved up into separate institutes.

Declan Butler

France's CNRS, the largest fundamental science agency in Europe, is to be reorganized into six quasi-autonomous national institutes by the end of the year. In essence, the move amounts to a dismantling of the CNRS, commentators say, replacing it with a UK-style system, which is organized by major discipline.

The current CNRS-controlled departments will be hived off into a federation of six new semi-autonomous national institutes for mathematics, physics, chemistry, engineering sciences, humanities and social sciences, and ecology and biodiversity. These will be similar in nature to its two existing institutes, the National Institute of Nuclear and Particle Physics and the National Institute of Earth Sciences and Astronomy.

Life sciences — which currently accounts for around one-quarter of the CNRS's budget — is strikingly absent from the list. National coordination of life-sciences research will transfer to the national biomedical agency Inserm, although the CNRS will still have a say, as will the agricultural research agency INRA and the atomic energy commission (the CEA). The national strategy for information-technology research will be shared between the CNRS and the national computer science agency INRIA. Both disciplines will be relegated to departments within the CNRS.

Another radical change is that the CNRS will not appoint the directors of the new institutes. Instead, it will be able to propose names, but the government will appoint the heads after an international search. According to Valérie Pécresse, the minister for higher education and research, this move will improve transparency and “attract the best scientists”.

The CNRS was due to announce its reform plans on 19 June after negotiations with the science ministry and other agencies. However, in a deliberate *fait accompli* on 20 May, Pécresse laid out the reforms in *Le Monde* newspaper. That tactic has been slammed by the science and higher-education trade unions and the ‘Save Research’ movement, who denounced the reforms as “asset-stripping” and a bid by the government

to seize research control from the CNRS. As Nature went to press, opponents of the reform were organizing 'academic pride' demonstrations across the country.

The oldest pregnant mum

Researchers in Australia have uncovered the oldest record of live birth — viviparity — in any vertebrate (see page 650). The discovery of embryos in fossils of placoderms (ancient, armoured, jawed fish) indicates that vertebrates have been copulating and giving birth to live young for at least 380 million years.

"We've discovered the world's oldest [pregnant] mother," says palaeontologist John Long at Museum Victoria in Melbourne, Australia, who led the study. Previously, the earliest records of viviparity were from marine reptiles from the Jurassic period that date back to around 180 million years ago, he says.

The fossils are from the Gogo Formation, a one-time coral reef in northwest Australia that is renowned for its remarkably well-preserved fish from the Devonian period, some 380 million years ago. The region is tectonically stable, so specimens have been spared the movements of Earth's plates that often distort fossils. "Gogo fish are three-dimensional, uncrushed, perfect specimens — as if they

sharks that are used to inseminate females.

"Internal fertilization does not have to mean viviparity, because some sharks and rays lay their eggs after they have been fertilized internally, but with the discovery of the embryos we can now be sure that at least some placoderms gave birth to live young," says palaeontologist Per Ahlberg at Uppsala University in Sweden.

Most modern fish opt for high-volume procreation, spawning masses of eggs, few of which survive. Placoderms, however, seem to have a similar reproductive strategy to that of some modern sharks and rays, investing in a few progeny and rearing them internally before releasing them as well-developed juveniles.

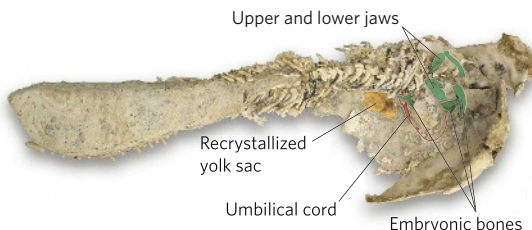
Placoderms are the first branch of jawed vertebrates and are thought to be a sister group of the common ancestor to both sharks and bony fish — viviparity is thought to have evolved independently in these different vertebrate lineages. The new discovery is important, Ahlberg says, because placoderms are completely extinct, unlike other evolutionary branches that have living descendants from which functional and structural information can be inferred.

"Our ongoing investigations suggest that viviparity is more widespread than previously thought in early placoderms," says Long, alluding to current efforts to scan other Gogo specimens for embryos. Indeed, one of the Gogo species from the current paper was first described by Long two decades earlier, but he had mistaken the embryos for an unusual cluster of external scales. "In earlier studies of placoderms, remnants of young have been observed inside adults and it was assumed to be cannibalism. Now we will have to look more carefully," says palaeontologist Philippe Janvier of the National Museum of Natural History in Paris.

Michael Coates, a palaeontologist at the University of Chicago, Illinois, predicts that there will be some debate over the soft-tissue preservation and whether the presumed yolk sac is a true structure. "It raises questions over the range of possible maternal-feeding strategies," Coates adds. Sharks use various strategies for supplying nutrients to their embryos, including secreting nutrients from the lining of the oviducts, or releasing unfertilized yolk-containing eggs for the embryos to consume.

Sharks have an elaborate and violent courtship, and these findings raise titillating speculation over placoderm sexual behaviour. "These early primitive fish, which were thought to be big, slow, dull, armoured fish, probably had an amazing courtship ritual," says Long. ■

Carina Dennis



A placoderm fish containing an embryo.

died yesterday," says Long. Soft tissues, including muscle and nerve structures, have been reported in Gogo specimens (K. Trinajstić *et al. Biol. Lett.* 3, 197–200; 2007).

The researchers identified a single embryo in a new Gogo fish genus, and three embryos in a previously described specimen. "When you find a little fish inside a big fish, you tend to think it was dinner," Long says. But the researchers concluded that the bones were those of embryos, not ingested remains, because they were not crushed or etched by digestive acids. What nailed it, according to Long, was the identification of an umbilical structure and a putative yolk sac.

Palaeontologists will be excited but not completely surprised by the findings, because many suspected that some placoderms fertilized internally. The males of a sub-group of placoderms, called *ptyctodontids*, have clasper-like appendages dangling from their pelvic fins — these are reminiscent of the claspers of modern

ON THE RECORD

"He's the best when it comes to dog cloning, and for that reason it behoves us to work with him."

Lou Hawthorne, chief executive of BioArts International in Mill Valley, California, explains why his company has partnered with disgraced South Korean scientist Woo Suk Hwang to clone pet canines. The bidding for the first five dog duplications will start next month at US\$100,000.

ROBOT NEWS

Robobunny

Researchers on two continents have made major advances towards developing robotic rabbits. In British Columbia, a scientist has built a therapeutic bunny that responds to petting. Meanwhile, a Swiss team has developed a microbot that hops hither and thither. Next step: designing an automaton fuelled by lettuce.



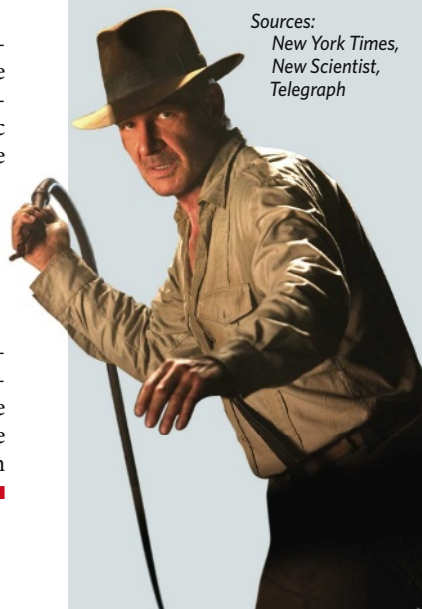
S. YOHANAN

SHOWBIZ NEWS

Indiana Jones and the wax crusade

Harrison Ford feels the planet's pain. Literally. The 65-year-old actor had his chest waxed to highlight how deforestation hurts the Earth's environment. Ford has had a busy month: he has been elected to the board of the Archaeological Institute of America and apparently also has a new film out.

Sources:
New York Times,
New Scientist,
Telegraph



D. JAMES (C) 2008 LUCASFILM LTD

Phoenix prepares for work after safe landing on Mars

After a dramatic deceleration through Mars's thin atmosphere, NASA's Phoenix mission is set to go to work. It landed on 26 May about 30 kilometres from its intended set-down point in Mars's northern plains. Its arrival marks the first successful 'soft' landing on Mars — a gentle setting down by means of retrorockets, as opposed to a crash-bang-wallop arrival swaddled in airbags — since that of Viking 2, more than 30 years ago.

"This is a scientist's dream, right there on this landing site," says mission principal investigator Peter Smith of the University of Arizona at Tucson, looking at pictures that show distinctive polygonal patterns on the surface (see right). Such patterning is one of the lines of evidence suggesting that Phoenix's digging arm should be able to unearth ice from the subsurface. The ice may record changes in the martian climate, and could contain organic molecules.

Mission scientists will spend a few days checking and calibrating instruments before starting to take samples for mass spectrometry and chemical analysis.

The successful descent was documented by a phenomenal image of the lander's parachute and shell (see inset), taken by the HiRISE camera on NASA's Mars Reconnaissance Orbiter during the descent from a distance of around 760 kilometres.

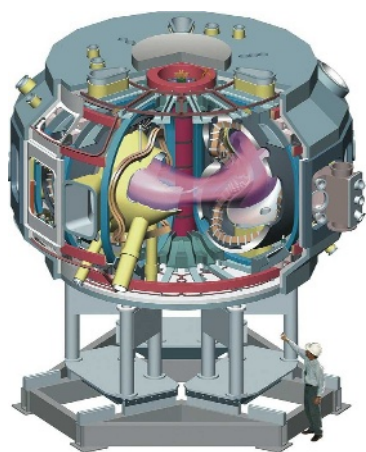


Norway's first Kavlis take on Nobel prizes

Research on quasars and carbon nanotubes was recognized by the judges of the inaugural Kavli Prizes, announced on 28 May in Oslo. Maarten Schmidt of the California Institute of Technology, and Donald Lynden-Bell of the University of Cambridge in the UK, are joint recipients of the astrophysics prize for their pioneering work on quasars. This "dramatically expanded the scale of the observable Universe and led to our present view of the violent Universe in which massive black holes play a key role," according to the Kavli committee.

Louis Brus of Columbia University in New York and Sumio Iijima of Meijo University in Nagoya, Japan, share the prize for nanoscience for their respective discoveries of colloidal semiconductor nanocrystals (quantum dots) and carbon nanotubes. And the neuroscience prize — awarded for their "discoveries on the developmental and functional logic of neuronal circuits" — went to Pasko Rakic of Yale University School of Medicine in New Haven, Connecticut, Thomas Jessell of Columbia University and Sten Grillner of the Karolinska Institute in Stockholm. The three US\$1-million prizes are to be awarded every two years.

PPPL



The National Compact Stellarator Experiment never made it out of the design phase.

Plug pulled on stellarator fusion project

The US Department of Energy has cancelled a major fusion experiment.

Based at Princeton Plasma Physics Laboratory in New Jersey, the US\$100-million National Compact Stellarator Experiment (NCSX) faced construction delays and cost overruns (see *Nature* 449, 264; 2007) estimated in terms of years and tens

of millions of dollars.

On 22 May, Raymond Orbach, the department's undersecretary for science, announced that NCSX would be terminated to free up funds for other fusion experiments, including ITER, an international \$12-billion tokamak under construction in Cadarache, France.

Funding boost for B-cell-based HIV vaccine research

In an effort to speed up HIV vaccine research, the US National Institute of Allergy and Infectious Diseases (NIAID) last week awarded US\$15.6 million in grants for research on B-cell immunology.

The five-year awards will go to ten research teams. They come two months after NIAID director Anthony Fauci announced his institute's intention to "turn the knob" in the direction of basic research and discovery (see *Nature* 452, 503; 2008) after the highly public failure of a T-cell-based HIV vaccine candidate.

"This is the kind of thing we were talking about when we were talking about discovery," says Fauci, who says that planning for the awards began 14 months ago. B cells make antibodies that neutralize invading viruses, but humans

seem unable to mount a response during HIV infection.

Japan to allow limited human embryonic cloning

Last week, an expert committee in Japan's science ministry agreed to lift a 2001 ban on human cloning for research purposes. The amendment, expected to be effective later this year, maps out rigorous ethical regulations under which the research could proceed. Reproductive cloning would still be illegal.

A group in Japan has expressed intentions to clone embryos for research. But serious hurdles remain, including the low availability of human eggs for research due to strict prohibitions on compensation.

Correction

The 2002 biosolids study from the National Academy of Sciences (NAS) did not reference research into health impacts by Environmental Protection Agency (EPA) whistleblower David Lewis, as reported in our News story 'Raking through sludge exposes a stink' (*Nature* 453, 262–263; 2008). The citation was included in a prepublication draft that is still posted on the EPA's internet site, but the NAS panel voted to remove the reference before final publication. An NAS spokesman said the panel decided the information was not relevant as the panel was not charged with evaluating health impacts.

The human body teems with microbes. In the first of two features, **Asher Mullard** looks at the global efforts to catalogue this vast 'microbiome'. In the second, **Apoorva Mandavilli** meets the surgeons who have a rare opportunity to watch an ecosystem being established as they transplant guts from one person to another.

THE INSIDE STORY

Any story about a human's microbes tends to invoke impressive numbers. Take the 10 trillion or so microbial cells living in the gut, which exceed the number of human cells by 10 to 1. Between them, they harbour millions of genes, compared with the paltry 20,000 estimated in the human genome. To say that you are outnumbered is a massive understatement.

But that might not be a bad thing. There is strength in numbers; so much so, in fact, that some biologists regard a human as a 'super-organism' — a community that adds up to more than the sum of its parts. The body itself is merely one, albeit encompassing, component.

Some smaller but nonetheless striking numbers about human microbes come in cash amounts. Late last year, the US National Institutes of Health (NIH) pledged US\$115 million to identify and characterize the human microbiome, the name given to the collection of microorganisms living in and on the human body. Also last year, the European Commission and various research institutes committed €20 million (US\$31 million) to similar ends. And smaller sums are being thrown in by funding agencies in countries that include

China, Canada, Japan, Singapore and Australia (see map).

Given the multifaceted nature of the microbiome, perhaps it is only right that it is studied by a large and varied community. The NIH's five-year Human Microbiome Project will spend much of its money identifying which bacteria are lodged where in the body and compiling a reference set of their genetic sequences. By contrast, the European Commission's four-year initiative, called Metagenomics of the Human Intestinal Tract (MetaHIT), will focus on microbial inhabitants of the gut, the main repository of the microbiota, and how they contribute to obesity and inflammatory bowel disease. Researchers involved in these and other initiatives say they will team up within a larger international consortium, but hints of competition simmer beneath the surface. "The intention is to work together," says George Weinstock, a geneticist at Washington University in St Louis, who is helping to organize the Human Microbiome Project, "but for the moment it is more about working in parallel until we can understand how to work together".

The microbes that swarm in and on the

human body have always held a certain fascination for researchers. Studies over the past century have shown that mice raised in a germ-free bubble have weak immune systems, inefficient digestive systems and abnormally small internal organs. They have shown that microbes are also an essential part of human biology.

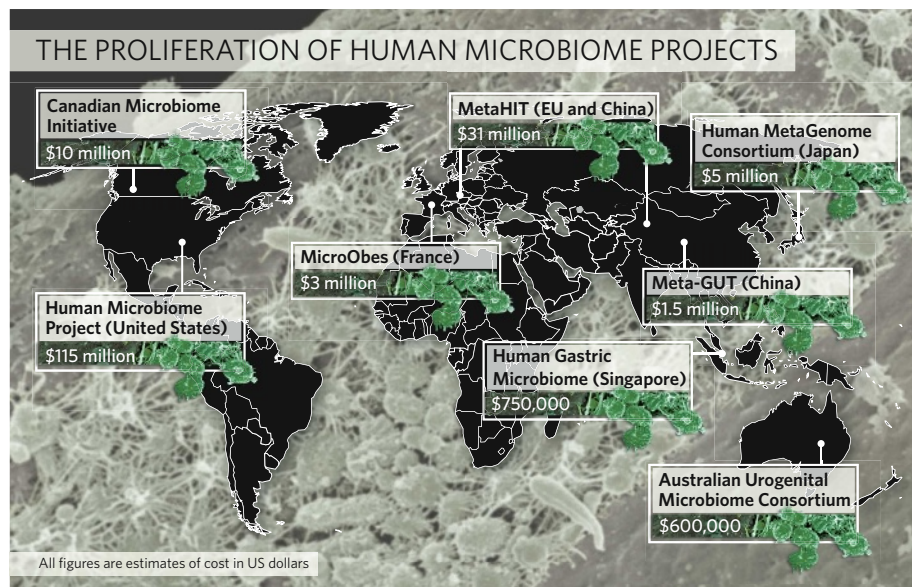
For the germ-ridden human masses who don't live in plastic bubbles, however, it has been difficult to work out exactly who these microbial passengers are and how they interact with one another. This is partly because there are so many of them, and partly because so few are easily grown in the lab.

A numbers game

Things have changed in the past few years. A few million foreign genes no longer sound so daunting in the face of advanced genetic-sequencing methods that have staggering numerical endowments of their own. Using the newest technologies, researchers can process hundreds of millions of base pairs in just a few hours. So they can bypass the need to grow bacteria, exploring the human microbiome by studying genes en masse, rather than studying the organisms themselves.

In 2006, Steven Gill at the Institute for Genomic Research in Rockville, Maryland, and his colleagues, threw around some then-hefty numbers when they carried out such a metagenomic analysis of the microbes in two people's intestines. After 2,062 polymerase chain reactions and 78 million base pairs, the researchers still had only a sketch of the gut's genetics, but they had revealed an abundance of genes that are involved in producing amino acids and in other aspects of human metabolism¹. Although interest in large-scale studies of the human microbiome has been mounting over the past few years, it was the realization that sequencing technology really was up to the task that finally prompted the money to roll in, Weinstock says.

Preliminary surveys such as Gill's have indicated that genes, age, diet, lifestyle and geography all affect which bacteria live in a person's body. But these first surveys involved too few





Microbes in the human gut may offer a wealth of information about health and disease.

for all big sequencing efforts and in this case, one-quarter of the money has been earmarked to examine the role of the microbiota in health and disease. Weinstock says that the project's main goal is to build up a research community and to generate a sequence resource, akin to that developed during the Human Genome Project, so that other basic and applied questions can be tackled later on.

MetaHIT is a different sort of beast. It will concentrate almost exclusively on the role of the gut microbiota in obesity and inflammatory bowel disease. And whereas the Human Microbiome Project is initially comparing people's microbiota on a species level, MetaHIT aims to find differences in microbial genes and the proteins they express without necessarily worrying about which species they came from. "We don't care if the name of the bacteria is *Enterobacter* or *Salmonella*. We want to know if there is an enzyme producing carbohydrates, an enzyme producing gas or an enzyme degrading proteins," explains Francisco Guarner, a gastroenterologist at Vall D'Hebron University Hospital in Barcelona, Spain. We want to "examine associations between bacterial genes and human phenotypes", says Dusko Ehrlich, coordinator of MetaHIT and head of microbial genetics at INRA, the French agricultural research agency in Jouy-en-Josas.

Gut reactions

Jeffrey Gordon, a microbiologist at Washington University School of Medicine in St Louis, has already shown how illuminating these associations can be. Gordon, one of the pioneers in the human microbiome field, showed two years ago that obese and lean individuals have radically different profiles of bacteria in their guts². When the obese volunteers went on a one-year diet and lost up to one-quarter

of their bodyweight, their bacterial profiles changed to look more like those of the lean people. The theory, based on studies in mice³, is that part of the propensity to gain weight might lie in 'obesity-causing' bacteria in the gut that release more calories from food than those found in lean people.

Researchers hope to gain further insight into how this happens by comparing the microbial genes in thin and fat people, and the findings could help to determine whether probiotics or other interventions could be used to shape the microbiome.

Some of this work is already under way as part of MetaHIT. In Denmark, a team led by Oluf Pedersen at the Steno Diabetes Centre in Copenhagen is collecting faecal samples from

individuals and sampled too few microbes, usually from only the gut or the mouth, to provide an adequate description of the microbiome. How many bacterial species colonize the entire body remains anyone's guess. So does the question of which ones everyone shares. "One of the things that is obsessing microbiologists is: 'What is the size of the core microbiome,'" says Jeremy Nicholson, a biological chemist who studies microbes and metabolism at Imperial College London.

Bringing order to chaos

The Human Microbiome Project is just the project for such obsessed microbiologists. In this, its first year, researchers will collect samples of faeces plus swabs from the vagina, mouth, nose and skin from 250 volunteers. They will sequence short, variable stretches of DNA that code for components of ribosomes in order to roughly identify which bacteria are present in each person and how many the volunteers have in common. With an estimate of diversity in hand, the researchers then plan to mine deeper. For this, they will also use shotgun sequencing to analyse many short pieces

of DNA from all over the microbes' genomes and reveal which genes are present.

Like any biology project that involves large sums and grand aims, the Human Microbiome Project has brought out some critics. One challenge, they say, is that the core microbiome might be incredibly small. Even though two people may each have 1,000 types of bacteria living in their guts, they might have only 10 species in common, for instance. And the commonalities might not lie in the genes, but rather in bacteria that have the same metabolic and physiological role. The Human Microbiome Project will do little to assess the function of microbes during its first year, although it may focus on this later.

Sarkis Mazmanian, a microbiologist at the California Institute of Technology in Pasadena, voices another reservation. "There's very little in terms of actual application to disease as part of the initiative. The lion's share of the efforts is in sequencing." This criticism tends to be aired

"One of the things that is obsessing microbiologists is: what is the size of the core microbiome."
— Jeremy Nicholson

120 obese volunteers and 60 controls to tease out specific microbial genes that might contribute to obesity. A similar-sized study in Spain, led by Guarner, will compare the microbiotas of patients with inflammatory bowel disease with those of genetically matched controls and examine the effect of drugs.

Whatever their differences, those involved in the Human Microbiome Project and MetaHIT sometimes find themselves on common ground — and that comes down again to those mind-boggling numbers. How do you effectively study such a vast and unknown community? Both groups will be using shotgun sequencing, which generates scraps of sequence from the many different genes and species. To make the most of this approach, a reference sequence is needed against which researchers can compare and identify these scraps, but at present there aren't enough known bacterial gene sequences to match the fragments against. Earlier studies, such as Gill's, produced many sequences that were impossible to assign functions to.

Into the unknown

Both the Human Microbiome Project and MetaHIT plan to sequence the complete genomes of hundreds of bacterial species and deposit them in a shared database. They will initially use shotgun sequencing of a few select species that can already be grown, and piece together their whole genomes from overlapping fragments. The Human Microbiome Project plans to provide 600 'reference genomes', MetaHIT will do another 100, and other sequencing efforts by the NIH and elsewhere will make additional contributions. With a broad enough reference database, researchers hope to be able to predict the genetic capabilities of some of those recalcitrant, unculturable species solely on the basis of similarities with known genes.

Even with such a reference, "it is pretty hairy from a computational biology analysis point of view", says Peer Bork, the biochemist who heads MetaHIT's computational centre at the European Molecular Biology Laboratory in Heidelberg, Germany. Even with the immense power of supercomputers to process the sequencing data, it will take some clever analysis to compare the millions of sequence reads that span thousands of species between hundreds of healthy and unhealthy people. It may be even hairier if, as many suspect,

subtle genetic patterns are what is important in disease rather than the presence of a single gene or species.

In many ways, such bioinformatics brings the need for collaboration into stark relief. When all the projects are running at speed, reams of data will be generated worldwide. But because different groups are using different techniques to collect samples, extract DNA and annotate data, the data sets are difficult to compare, Bork says.

Enter the as-yet-unlaunched International Human Microbiome Consortium. Scientists from several international projects, including the Human Microbiome Project and MetaHIT, have been meeting since late 2005 to figure out how to collaborate on a range of issues such as the compatibility of data and which bacteria to sequence for the reference database. The group is already setting up infrastructure and "beginning to address the tough questions", says Weinstock. But it is too early to say how well it will grease the wheels of collaboration. Its official launch, scheduled for April, was postponed for six months to allow the NIH and the European Commission to overcome bureaucratic hurdles. Even so, optimism for the collaboration runs high, partly because its members can still pursue their own pet projects. "Talented people are doing what they think is the most important research to do, rather than being forced to do what somebody else has decided would be the best," says Ehrlich.

One of questions being addressed by the consortium is over intellectual property. As with other genomic projects, members of the consortium will be expected to release sequence data into the public domain as soon as they are generated. But this doesn't necessarily preclude disputes over intellectual property if, for instance, a particular bacterial gene proves to be a useful diagnostic marker for a disease. Another unresolved question is whether a laboratory can have one project that abides by the consortium's regulations, and another that doesn't. "There are grey areas, and I feel that



"There's so much to learn, so much we don't know and so many adventures."
— Jeffrey Gordon

until we have a test case, they will have to be watched very carefully," says Bhagirath Singh of the Canadian Institute of Health Research, who is helping to develop the Canadian Microbiome Initiative.

Participants from microbiome projects around the world say that they plan to sign up to the consortium. But the number of independent projects being launched speaks to at least some underlying competition between those involved, and some of the difficulties the group may face. In addition to differences in focus and scope, MetaHIT's operating budget is only a quarter the size of the Human Microbiome Project's. "This is giving

a huge advantage to the Americans," Guarner says. "They are going to be quicker and they have more equipment." But some members of MetaHIT feel that they actually have an edge because money for their project has already been distributed and data collection is under way, whereas the Human Microbiome Project will not announce many of its funding decisions until later this year. "We have an advantage already, we have a show on the road," says Willem de Vos, a microbiologist at Wageningen University in the Netherlands and a member of MetaHIT.

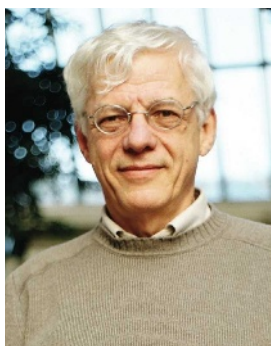
Given the number of separate projects, all at such an early stage, it's almost impossible to make out where the starting line lies or who, exactly, is edging ahead. And the intention to pool data means that there may be no clear line separating them. "If it is an international consortium, it doesn't matter where the data are generated," Bork adds. "For example, we can be the pirates here, sitting at the end in Europe, and use American data to make the discoveries."

Plundering may be unwarranted. With trillions of microbes to sift through, most researchers feel that there is more than enough of the microbiome to go around. "There's so much to learn, so much we don't know and so many adventures," Gordon says. "There's enough room for everyone."

Asher Mullard is an assistant editor for *Nature Reviews Microbiology* and *Nature Reviews Molecular Cell Biology*.

1. Gill, S. R. *et al. Science* **312**, 1355–1359 (2006).
2. Ley, R. E. *et al. Nature* **444**, 1022–1023 (2006).
3. Turnbaugh, P. J. *et al. Nature* **444**, 1027–1031 (2006).

See Editorial, page 563, and News Feature, page 581.



"Talented people are doing what they think is the most important research to do."

— Dusko Ehrlich

COURTESY OF WASHINGTON UNIV., ST LOUIS

J. MIGNOT

STRAIGHT FROM THE GUT

Stephanie is the first to admit that she never had the guts for life. She was born with familial adenomatous polyposis, a genetic disorder in which thousands of polyps form in the colon. By the age of 22, much of the organ had to be removed. Four years later, a massive benign tumour choked off the blood supply to her small intestine, so doctors cut out all but a metre of it. For the next six years, she was fed by a tube every night until the feeding left her liver badly scarred and fighting recurring infections. "I was given a month to live," she says.

That's when doctors referred Stephanie to Georgetown University Hospital in Washington DC. There, on 17 April 2006, surgeons cut out her stomach and what was left of her small and large intestine and replaced it with new organs from a donor who had died days earlier in Tennessee. "Oesophagus to anus, her entire gastrointestinal tract was in the garbage can," says Tom Fishbein, who directed the surgery. "She got a brand new one."

All organ transplants are complicated, but there are only a handful of centres in the United States that have the expertise to transplant a small intestine, the

seven metres of coiled tissue connected up to the stomach at one end and the large intestine at the other. The technique is complicated because the gut is teeming with trillions of bacteria and other microbes, plus the bulk of the body's lymphocytes. Before such transplants, the donor's intestine has normally been flushed with antibiotics. But rates of infection and rejection from such transplants are very high because, it is thought, some foreign bacteria and immune cells survive the cleaning process and are thrust into an immunosuppressed recipient.

The idea that these intestinal bacteria are a menace is now under review. By teaming up with microbiologists, the surgeons are taking advantage of a rare chance to study microbes as they colonize the walls of the gut after transplanting an intestine: which ones arrive first, and how they restore the ravaged microbial communities. "An all new ecosystem of organisms had to populate that bowel from scratch," says Fishbein. Their new appreciation of that ecosystem, along with their growing surgical experience, suggests that the populations might be better left intact before a transplant.

The same studies may also offer insight into

how the gut is first populated by microbes after birth, how it recovers from the damage done by a heavy course of antibiotics and, perhaps, how to minimize that damage. "Most people study this in animal models, but this is a real-person model," says Brett Finlay, a microbiologist at the University of British Columbia in Vancouver, Canada. "It's an artificial system in some sense, but it's a neat model."

"Oesophagus to anus, her entire gastrointestinal tract was in the garbage can."

— Tom Fishbein

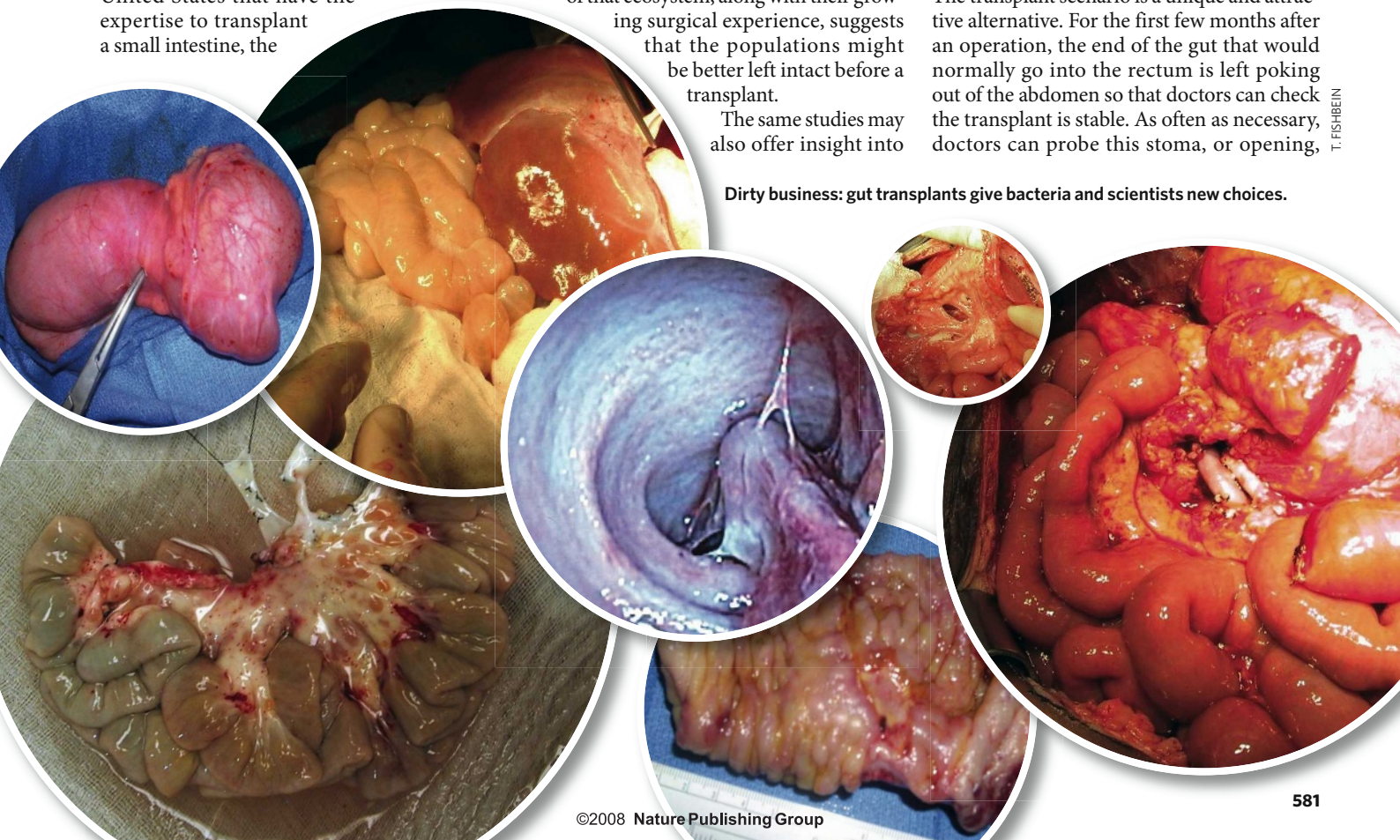
How the moist, pink intestinal tubing lives in such harmonious contact with bacteria has puzzled scientists for decades. But "it's hard to get in there, especially in a healthy person", says David Relman, who studies microbiology and immunology

at Stanford University, California. "And to do it in a way that doesn't perturb the system, and to do it every week or every day, well, forget it." For this reason, most researchers interested in the contents of human innards have had to collect and filter faeces.

Beautiful opportunity

The transplant scenario is a unique and attractive alternative. For the first few months after an operation, the end of the gut that would normally go into the rectum is left poking out of the abdomen so that doctors can check the transplant is stable. As often as necessary, doctors can probe this stoma, or opening,

Dirty business: gut transplants give bacteria and scientists new choices.



T. FISHBEIN



An opportunity to track gut colonization.

with an endoscope to pinch off biopsies of the intestinal wall with its microbial community intact. Yellow faeces and glistening, pink skin are signs that the new intestine has successfully taken. "That is beautiful stuff," says Stuart Kaufman, medical director of the intestinal-transplantation programme at Georgetown. "We live for poop like this."

The adult samples are also beautiful stuff for researchers who want to chart the arrivals of the different bacterial species. The material is put on ice and shipped to the lab of Jonathan Eisen, a microbial geneticist at the University of California, Davis. There, graduate student Amber Hartman identifies the inhabitants by analysing distinguishing gene sequences that vary slightly between different species of bacteria.

It's far too early to draw firm conclusions from their data; Fishbein and his collaborators have only studied 15 patients over the past 2 years. But what they have found suggests that the gut is populated first by enterobacteriaceae, a large family of facultative anaerobes, which can grow with or without oxygen. This suggests that the transplanted tissue has higher oxygen levels than the normal gut, where anaerobic bacteria dominate. Inflammation may boost oxygen levels, giving enterobacteria the advantage.

Hartman and Eisen have also found that each person studied so far has had different proportions of the various microbial species, and that these oscillate rapidly in the first few months after a transplant. Their preliminary observations suggest that the more chaotic the variations over time, the worse the outcome of the transplant. "One thing that becomes very obvious is that the amount or degree of fluctuation is much greater in the sicker patient," Hartman says. They now want to know whether the microbiota becomes more stable and reaches equilibrium as a patient recovers, but that

requires following more patients over longer times, and intestinal transplants are rare.

A different bloody, messy procedure arrives all too often though: a birth. Here, too, researchers see a fascinating opportunity to explore how microbes colonize a gut, one thought to be sterile inside the womb. During birth and in the hours after, babies can swallow bacteria from the mother's birth canal, faeces and from whatever environment they arrive in.

No one knows yet whether bacteria move into a baby's spanking new innards in the same way they grab a foothold in a used, adult transplant. Much of what scientists know about the former process has been learned from a study of Relman's, in which he and his colleagues collected used nappies from 14 babies beginning with the first stool after birth and at regular intervals throughout the first year of life¹. What they found mirrored some of the discoveries in the transplanted intestines: every baby's microbiota is unique but dynamic, with different populations of bacterial species shifting in abundance. And as in the transplant study, the babies showed a succession of colonization, with facultative bacteria settling in first, followed by a more complex and diverse population.

Nice and dirty

Whatever parallels may emerge from these studies, there is one obvious difference: the transplanted gut has already been soiled by the faeces, microbes and immune cells it hosted before. And, until recently, doctors did all they could to scour away the muck.

Paradoxically, the surgeons at Georgetown began to notice that the more antibiotics they used to keep the microbes to a minimum, the more intestinal infections they saw after the transplant. At first, the alternative seemed too fantastic to contemplate: taking an essentially infected organ and placing it in a body crippled by immunosuppressant drugs. But about a year ago, Georgetown and other centres began shifting their practice to do exactly that. Early evidence indicates that those who receive a gut replete with its native microbiota have fewer chaotic fluctuations.

That makes sense in retrospect, notes Eisen, because people who need transplants may do so precisely because they had trouble colonizing their bowels properly to begin with. In some patients — those with Crohn's disease or ulcerative colitis, for example — the native bacteria may turn on the body, inflaming the

gut and scarring the intestinal walls. So the new intestine might be more likely to help if it comes with its own set of inhabitants. "To take out all the microbes seems completely inane," Eisen says.

The Georgetown researchers have now started to investigate how those bacterial fluctuations during colonization are controlled. They suspect that a gene called *NOD2*, which is expressed in some immune cells, is essential for keeping the chaos to a minimum. The *NOD2* protein recognizes components of bacterial cell walls and controls the production of defensins, small proteins that kill particular bacterial species.

The team has found that about 35% of their transplant patients carry mutations in *NOD2*, regardless of the intestinal disorder, and that those with mutations are 100-fold more likely to have a failed transplant compared with controls². The hypothesis is that this mutation somehow lowers production of defensins, so the immune system is unable to maintain the appropriate proportions of bacterial species. Perhaps patients with a mutant *NOD2* gene might benefit from doses of the bacteria that they are missing. In one patient with a *NOD2* mutation who later died, "the proportions of normalness were very, very off", Hartman says.

Quite what 'off' is, is hard to define. The team still has only a cursory understanding of what the microbiota looks like in healthy people, compared with their subjects. "It's like watching colonization of a disturbed ecosystem without knowing what was originally in the forest," says Eisen. That may be helped by the National Institutes of Health's Human Microbiome Project and other new research efforts that aim to catalogue the microbes in the human body (see page 578). "We need that field guide to microbes to understand when something is not normal," Eisen says.

In the meantime, clinical signs are still the best predictors of a transplant's success. Two years after a new gut was slotted into Stephanie's body, she still has scars crisscrossing her abdomen but she is a healthy weight and eats whatever she wants. It is still not clear which microbes to thank: the donor's bugs that survived from before the transplantation, or new microbes that settled there afterwards. "Whoever's microbes have prevailed, they're probably good ones," says Fishbein, "because she's done exceptionally well."

Apoorva Mandavilli is a freelance writer based in New York.

1. Palmer, C. et al. *PLoS Biol.* 5, e177 (2007).

2. Fishbein, T. et al. *Gut* 57, 323–330 (2008).

See Editorial, page 563, and News Feature, page 578.

CAMERON/CORBIS



Origins of DEATH

Programmed cell death is usually seen as the unique prerogative of plants and animals. So how is it that photosynthetic plankton have been killing themselves by uncannily similar methods for billions of years? **Nick Lane** investigates.

One evening 20 years ago, Paul Falkowski left the lab so tired that he omitted to refresh the solution in his culture flasks of *Emiliania huxleyi*, one of the world's most widespread coccolithophores. The following morning he was shocked to find the flasks full of clear solution, the merest sediment lining the bottom, all that remained of the plankton. "I had never seen anything like it," he recalls. "They just dissolved overnight."

The speed and totality of their demise couldn't be put down to run-of-the-mill mortality, which leaves a gory mess of living cells, dying cells and clumps of dead matter. The tidy dissolution of the plankton was more reminiscent of apoptosis in animals — a synchronized wave of death orchestrated by some invisible hand to some unknowable end. Falkowski became determined to reveal that hand and understand its purpose, and now, in his lab and others, the revelations are beginning to add up.

Now a professor of marine biogeochemistry at Rutgers University in New Jersey, Falkowski is better known for his research on marine productivity and nutrient cycles. From that mainstream perspective, his interest in programmed cell death among plankton might seem a quirky morbid streak.

Dying, though, is something phytoplankton do a great deal of. Phytoplankton — bacteria

and eukaryotes that photosynthesize, such as coccolithophores and diatoms — fix as much carbon every year as all the plants on all the continents. Yet at any one time they account for just 1% of Earth's biomass. This means their rate of turnover is huge; on average, the world's phytoplankton population is replaced once a week.

Most models of marine systems simply put that mortality down to bad luck. They tacitly assume that phytoplankton are in principle immortal, but in practice always seem to be eaten by zooplankton, wiped out by viral infections or starved by nutrient deprivation. But Falkowski thinks that there is more to it than that. Phytoplankton don't just dissolve when neglected in the lab. Vast marine blooms, too, can disappear overnight. If that process is indeed a manifestation of programmed cell death, then it has implications that not only touch on the global nutrient cycling, but also foreshadow the ultimate causes of our own mortality. The phytoplankton of the ancient oceans might have been some of the first creatures to learn how to die.

Falkowski's hunch that the dissolution of his phytoplankton might have parallels with apoptosis in animals has since turned out to

be remarkably close to the mark. Apoptosis is masterminded with extraordinary finesse by a group of protein-splitting enzymes known as caspases. The scissor-like ability to cut proteins that sets these enzymes apart is constrained to very particular sequences of amino acids. They operate in cascades in which one caspase activates the next by slicing through a protective protein sheath. Each step amplifies the death signal until an army of executioners has been let loose on the cell, dismantling membrane structures, slicing up DNA, dicing proteins.

This carefully orchestrated cascade leaves tell-tale signs at the microscopic level. The nuclear material condenses; chromosomes fragment; the cell membrane blebs off into tiny bubbles; internal organelles disintegrate; then the cell dissolves around them.

Working with Kay Bidle, a marine microbiologist also at Rutgers, Falkowski found that cultured phytoplankton precisely reproduced every microscopic step. The pair also found that antibodies raised against human caspases cross-reacted with extracts taken from phytoplankton, suggesting that the extracts must contain proteins similar to the caspases¹. Furthermore, molecules known to be specific

"Much is owed to a lucky choice of bacteria." — Eugene Koonin

C. DARKIN

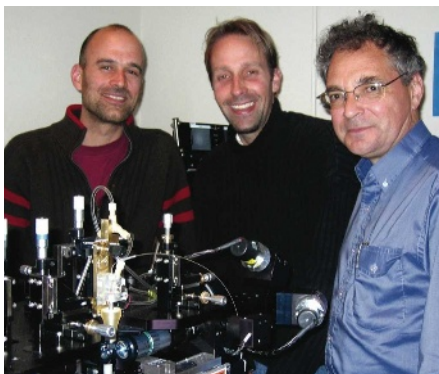
inhibitors of caspases interfered with the cells' death programme. Remarkably, the evidence of programmed death was not constrained to eukaryotes; it was found in photosynthetic cyanobacteria too.

By the late 1990s, similar findings had been reported in plants. But no one had managed to identify the caspases presumed responsible, either in the plants or in the plankton. Then in 2000, researchers at the National Center for Biotechnology Information in Bethesda, Maryland, and the biotechnology firm Genentech in San Francisco, California, discovered a related family of proteins by analysing genes from a wide range of species for sequences that might produce caspase-like activity when expressed². The *in silico* discovery of these 'metacaspases' paved the way to understanding programmed cell death in plants, algae and fungi, and offered the first real insight into the origin of the caspases, now known to be found only in animals.

Cosmopolitan lifestyle

Within a couple of years, Eugene Koonin and Lakshminarayanan Aravind, two of the researchers on the original discovery team, had sketched out the protein family's distribution across the tree of life³. In doing so, they confirmed the molecules' presence within a few groups of bacteria that show a touch more complexity than most of their peers, notably the photosynthetic cyanobacteria, which sometimes form multicellular chains. It was already clear, Koonin recalls, that the metacaspases and caspases derived from the same ancestor. Identifying that ancestor, though, is hard. Bacteria have very fluid chromosomes, and frequently incorporate genes derived from totally unrelated bacteria. So although cyanobacteria have the richest metacaspase reservoir, and thus might seem a likely point of origin, two other groups of bacteria have metacaspase genes and so need to be considered: the actinomycetes (a group of soil bacteria) and a few of the α -proteobacteria.

The most compelling account of the origin of mitochondria, the 'power plants' of almost all modern eukaryotic cells, is that they are descended from the α -proteobacteria. Metacaspases are widespread among eukaryotes, so they were probably acquired at an early point in eukaryotic history. "The fact that some α -proteobacteria contain metacaspases makes mitochondria the most likely source of ancestral metacaspases in eukaryotic cells," Koonin says. How the α -proteobacteria got their metacaspases, though, is an open question, and transfer from cyanobacteria is utterly possible.



Sorting out cells: Assaf Vardi (left), Kay Bidle (centre) and Paul Falkowski are trying to decipher why phytoplankton might want to kill themselves.

It is also possible that some eukaryotes received their metacaspases from two different lineages — once via the α -proteobacteria that became the mitochondria, and once via the engulfed cyanobacteria that evolved into chloroplasts. Bidle points out that some of the metacaspases found only in algae and plants are aimed at targets in chloroplasts, which hints they may have come in that way.

But why do cyanobacteria have all those metacaspases? In their 2002 paper³, Koonin and Aravind argued that, in general, bacterial metacaspases have some sort of signalling role; but they also wondered whether, in a few more-complex 'cosmopolitan' bacteria, they might mediate cell death as the animals' caspases do. And that was exactly what Bidle and Falkowski were wondering, too.

Metacaspases are in at least one sense not caspases: for one thing few, perhaps even none, cleave the classic caspase target sequence. For this and other reasons some researchers, notably Frank Van Breusegem and his colleagues at Ghent University in Belgium, doubt that metacaspases execute cell death in plants and phytoplankton — or at least whether the case is yet properly proved.

Bidle and Ilana Berman-Frank, at Bar Ilan University in Israel, have put some effort into that proof. They have shown that the number of metacaspases rises and that caspase activity sky-rockets during cell death in cyanobacteria such as *Trichodesmium* — a subtropical bloom-forming bacterium that might have evolved as long as three billion years ago. The microscopic demise of these cells looks like apoptosis — and the cyanobacteria in question turn out also to make several other key enzymes found in the cell-death cascade of animals. (Apoptosis-activating factor, apoptotic ATPases and

NACHT-family NTPases, since you ask.)

But this evidence is circumstantial, and the same team has also found a lot of metacaspases that are not involved in cell death. This February, for example, Bidle and his colleagues reported that *Thalassiosira pseudonana*, a single-celled diatom, has six putative metacaspases⁴, not far short of the nine found in plants such as the thale cress *Arabidopsis*, or indeed the twelve caspases found in humans. Only two of these are activated during cell death, however. The rest are apparently expressed in healthy cells, or sometimes in stressed ones, as found by Koonin. "Although it looks as though some metacaspases really are involved in cell death, they certainly seem to have biological functions that go beyond apoptosis," says Bidle.

The clearest-cut evidence for metacaspases in cell death was published this January by Patrick Gallois from the University of Manchester, UK, and his co-workers, who showed through careful genetic engineering of *Arabidopsis* that at least one metacaspase does directly mediate cell death⁵. In light of these findings, as well as some older reports from yeast and phytoplankton, Gallois thinks that metacaspases are probably "part of an ancestral pathway that activates programmed cell death in response to a damaging level of oxidative stress".

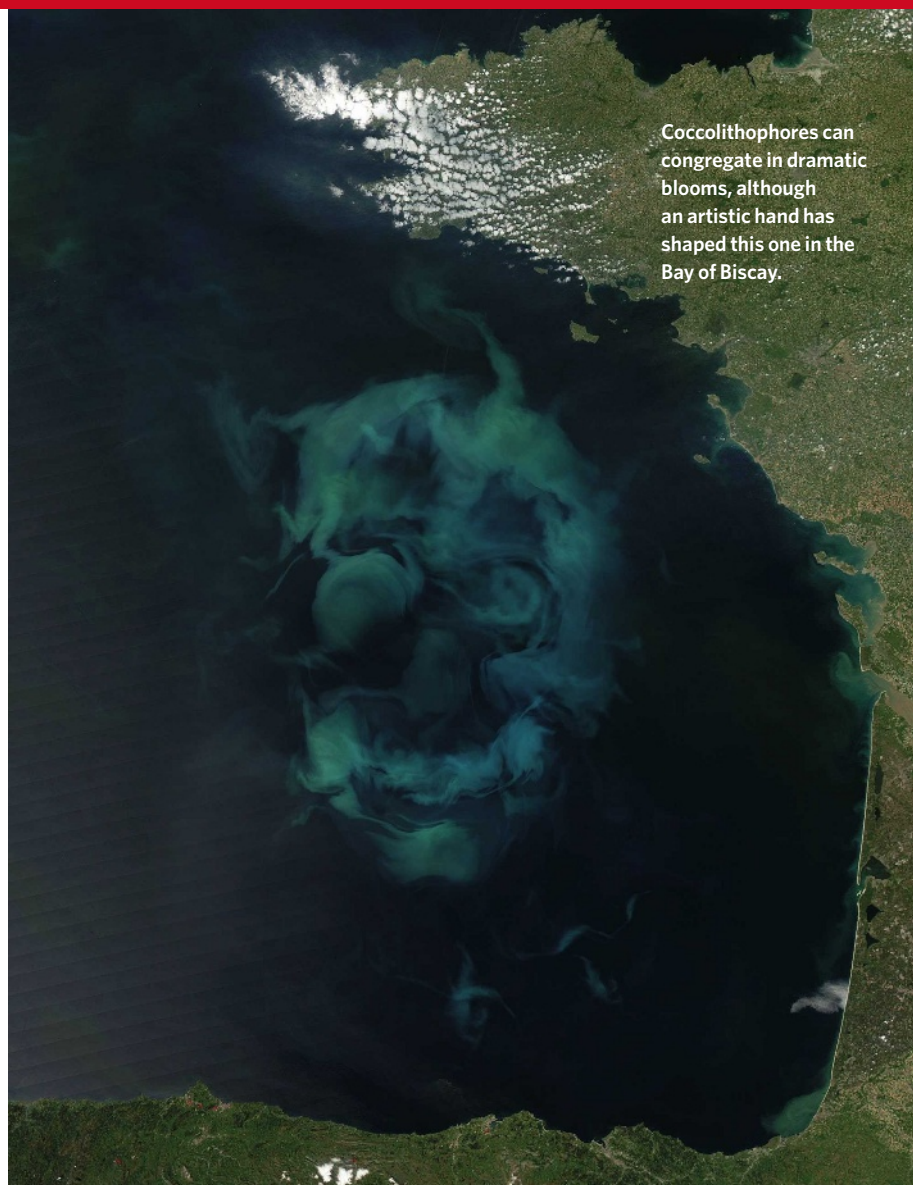
Stress relief

The tie-in with oxidative stress — a state in which there is more reactive oxygen around than a cell can tolerate — emphasizes another strong link with phytoplankton, says Assaf Vardi, a molecular biologist who joined Falkowski and Bidle at Rutgers in 2007. Ultra-violet radiation, carbon dioxide limitation, iron deficiency and viral infection can all trigger cell death in eukaryotic algae, he says. And what these things have in common are reactive oxygen species — superoxides, hydroxyls and the like. Long perceived merely as pathological by-products, these reactive oxygen species are now known to be used by cells as signals.

Vardi has found that in phytoplankton, signals from reactive oxygen species are amplified by nitric oxide. This gas does a great deal of signalling work in plants and animals, and has a key role in inflammation, immunity and cell death. Vardi is the first to show that the gas is actively generated by phytoplankton, and he has even filmed it using fluorescence microscopy.

Vardi's latest work shows that *Phaeodactylum tricornutum* — a diatom, like *T. pseudonana* — calls on an enzyme found only in chloroplasts to generate nitric oxide. Cells that overproduce that enzyme grow slowly and are highly sensitive to stress, activating as many as six metacaspases during the cell-death programme. The steps involved are eerily familiar to those who study

"The plankton just dissolved overnight." — Paul Falkowski



Coccolithophores can congregate in dramatic blooms, although an artistic hand has shaped this one in the Bay of Biscay.

apoptosis. Nitric oxide blocks the electron-transport chains needed for photosynthesis and respiration; that generates reactive oxygen species, and they in turn trigger the death cascade.

This seems to show that Falkowski's hunch was correct. Stressed phytoplankton blooms generate nitric oxide, and reactive oxygen species then activate the machinery of death. This machinery includes (but is probably not limited to) metacaspases that are functionally related to the caspases found in apoptosis in animals. But why does so much of the cell-death apparatus associated with complex animals — in which cellular self-sacrifice can make obvious sense — operate in single-celled phytoplankton? What can a plankton possibly have to gain by killing itself?

Falkowski suspects that the ultimate brokers of death are viruses. Sea water contains viruses in shocking abundance — hundreds of millions of viruses per millilitre of sea water — and phytoplankton are the targets of many if not most of them. Falkowski thinks that the death apparatus is part of an ancient tug of war between viruses and their prey.

Falkowski draws an analogy to the 'addic-

tion modules' found in simple bacteria such as *Escherichia coli*. In that case, viral genes in the cell encode a long-lived toxin as well as its short-lived antidote; cells that stop expressing the viral genes run out of antidote and die.

Power struggles

Cell-death systems that use metacaspases might have evolved in plankton through similar processes, Falkowski says, with viruses turning metacaspases into dangerous weapons of subjugation. Host cells then evolve new systems to wrest the viruses' hands away from the trigger. The system might be seen as a coevolved product of the virus and the host, with each trying to control it. Last year, Bidle and Falkowski showed that viral replication can be blocked by inhibiting the death apparatus of the coccolithophore *E. huxleyi*, implying that the virus usually uses the system to kill the cell when death is to its advantage.

The odd thing is that sometimes the cells may want to turn the death programme on when the viruses want to keep it turned off. If cells can kill themselves more quickly than suits their viral invaders, they can thwart the viral spread.

As most plankton in a bloom are near identical genetically, from the perspective of their genes, a die-off that creates enough scorched earth to stop the viral advance can make sense.

To understand this as self-sacrifice, though, might be to oversimplify. "It is at least as much murder as suicide," says Vardi. "It definitely blurs the boundary between altruism and selfishness." Vardi has found that injured phytoplankton release mediators — he calls them 'infochemicals' — into their surroundings. In response, damaged cells overproduce reactive oxygen species, activating the death apparatus, whereas in healthier cells the signals lead to several processes designed to deal with stress: the formation of tough, long-lasting cysts, the creation of biofilms on nearby surfaces, even differentiation into separate sexes. The system is assigning genetically identical cells different fates — rather as developmental programmes do in multicellular organisms.

Death is, in a way, the simplest form of development — a binary choice. It may be that caspases and their cousins have been called on as managers of death time and again, in shaping tissues by killing cells or in weeding out sicker plankton to benefit the rest, because these are the settings in which cells with the same genes need to thin their numbers. The programmed-death function may have been inherited from the cyanobacteria along with the metacaspases used for it. But it is also possible that the cascade-friendly features of such molecules lend themselves so well to a self-destruct sequence that they get roped in whenever such a capability is called for.

Either way, the links from today's complex world back to the origins of death in a simpler one seem oddly fortunate. The bacterial partners in the symbioses that led to eukaryotes might not have had anything as useful as metacaspases in their make-up. They might have been getting by with some simpler death system, such as an addiction module, or no death system at all. If so, eukaryotes could have started out without the cell-death mechanisms that are so crucial to their development. As Koonin puts it: "Much of the glorious ascension to the ultimate complexity of higher plants and animals is owed to a lucky choice of bacteria with complicated differentiation processes as partners in the origin of the eukaryotic cell." The complex seeds of death made life what it is today. ■

Nick Lane is author of *Power, Sex, Suicide: Mitochondria and the Meaning of Life*.

1. Berman-Frank, I., Bidle, K. D., Haramaty, L. & Falkowski, P. G. *Limnol. Oceanogr.* **49**, 997-1005 (2004).
2. Uren, A. G. et al. *Mol. Cell* **6**, 961-967 (2000).
3. Koonin, E. V. & Aravind, L. *Cell Death Differ.* **9**, 394-404 (2002).
4. Bidle, K. D. & Bender, S. J. *Eukaryot. Cell* **7**, 223-236 (2008).
5. He, R. et al. *J. Biol. Chem.* **283**, 774-783 (2008).

J. SCHWALTZ/ MODIS RAPID RESPONSE TEAM/NASA/GSFC (MODIFIED BY C. DARKIN)

CORRESPONDENCE

Increasing use of stimulants warns of potential abuse

SIR — Your Editorial 'Defining natural' (*Nature* **452**, 665–666; 2008) indicates that cognitive-enhancing drugs have only mild effects similar to caffeine, and the News story 'Poll results: look who's doping' (*Nature* **452**, 674–675; 2008) highlights the general increase in their availability through Internet purchase and possibly diversion of prescriptions. By far the most frequently prescribed of these drugs — the stimulant medications amphetamine and methylphenidate — have abuse potential and therefore warrant particular concern.

The World Health Organization (WHO) tracks by country the yearly production, supply and consumption of psychotropic substances, including stimulant drugs. The most recent figures suggest that we should evaluate carefully whether diversion of methylphenidate and amphetamine may be increasing in the United States. From 1995 to 2006, consumption estimates by the WHO and prescriptions recorded by the healthcare-information company Verispan both increased linearly. Moreover, the 268.9% relative increase in the WHO estimates of consumption, from 4.66 to 17.9 defined daily doses per 1,000 population, far exceeded the relative increase in number of prescriptions per year, from 15,044,359 to 30,137,136 (100.3%).

The greater increase in consumption could be related to a shift in the age of individuals treated for attention-deficit hyperactivity disorder, as more adults are now seeking diagnosis and treatment. This could influence debate over whether some prescriptions for clinical treatment of adults are being appropriated instead for performance enhancement or

recreational purposes. History teaches us that either of these could escalate into misuse.

We have serious concerns about the dramatic increase in national consumption and the smaller, but still large, increase in annual prescriptions of stimulant medications. These rises should make us vigilant to avoid a potential repeat of past episodes of abuse.

James M. Swanson University of California, Irvine, Department of Pediatrics, Child Development Center, 19722 MacArthur Boulevard, Irvine, California 92612, USA
Nora D. Volkow National Institute on Drug Abuse, 6001 Executive Blvd, Room 5274, Bethesda, Maryland 20892, USA

Costa Rica's biotech project still on track for end of year

SIR — In your News story 'Costa Rican biotech centre in peril' (*Nature* **452**, 787; 2008), you claim that the European Union (EU) is threatening to withdraw its contribution to the National Centre for Biotechnological Innovation (CENIBiot) project. We would like to clarify that the project is going ahead — we have not requested an extension of the EU's deadline, nor has the EU refused to grant one.

We are currently receiving sealed bids to remodel the CENIBiot building and equipment. All funds will be used before 5 December this year, as required under the EU agreement. No appeals are permitted for failed bids. Objections are dealt with according to EU procedure, without delaying application of the contract — unless a clear irregularity arises during the process, in which case it is referred to the EU. We have made every effort to guarantee maximum transparency in the process and to prevent violations.

For each bid, a tribunal of

independent technical evaluators is appointed (the identity of the evaluators is kept confidential), along with a president and secretary for the evaluating committee. Representatives from the EU, the Ministry of Foreign Relations and the Ministry of Science and Technology attend as observers to reinforce the transparency of the process.

The Strategic Development Unit of CENIBiot is preparing contracts for economic study packages, updating biotechnological databases in Costa Rica and organizing visibility contracts and the exchange of experience and capabilities.

All this information is available from the ministry, is in accordance with the 2008 operation plan of the project and is paving the way for the identification of CENIBiot demonstration projects, on which we are now working.

Antonieta Corrales Communication Office, Ministry of Science and Technology, Costa Rica

A prime problem that even quantum computing can't solve

SIR — Your News Feature 'The dreamweaver's abacus' (*Nature* **452**, 803–805; 2008) seems to have been rather carried away by the exciting prospects on offer from the emergent techniques of quantum computing. One thing that even quantum computing can never accomplish is "factoring large prime numbers". It cannot even factor small ones.

The author is in august company, however, as no less a personage than Bill Gates has had similar visions. In connection with the need to break modern encryption, he wrote: "The obvious mathematical breakthrough would be development of an easy way to factor large prime numbers" (in *The Road Ahead*; Viking Press, New York, 1995).

Robert P. Bywater Magdalen College, High Street, Oxford OX1 4AU, UK

Analyses support theory of stochastic regulation of fisheries

SIR — The month that saw the passing of David Cushing, perhaps Britain's leading marine fisheries ecologist, also saw evidence in *Nature* for his little-known theory of stochastic regulation of fish stocks.

Colin Minto and his colleagues show that both the survival and variability of fish larvae increases as the size of the population declines to give a density-dependent regulation (C. Minto, R. A. Myers and W. Blanchard *Nature* **452**, 344–347; 2008). Christian Anderson and his co-workers further explain how changes in the demographics of populations can generate these effects, which are magnified by fishing (C. N. K. Anderson *et al.* *Nature* **452**, 835–839; 2008).

Almost 20 years ago, Cushing and John Shepherd speculated that increased variability in larval survival provides the strong regulation that allows intensive fishing, and demonstrated the effects through modelling (J. G. Shepherd and D. H. Cushing *Phil. Trans. R. Soc. Lond. B* **330**, 151–164; 1990).

At the time, the hypothesis was left as plausible, likely even, but unresolved, as examinations of data proved inconclusive. The recent metadata analyses in *Nature* provide support for more variability when the population is small and so confirm at least a role for stochastic regulation in fish stocks. The analyses also reveal little or no clear increase in variability in many fish stocks, including many cod stocks. Other regulatory mechanisms might also exist.

Joe Horwood Centre for Environment, Fisheries & Aquaculture Science, Department for Environment, Food and Rural Affairs, Lowestoft, NR33 0HT, UK

John Shepherd School of Ocean and Earth Science, National Oceanography Centre, University of Southampton, SO14 3ZH, UK

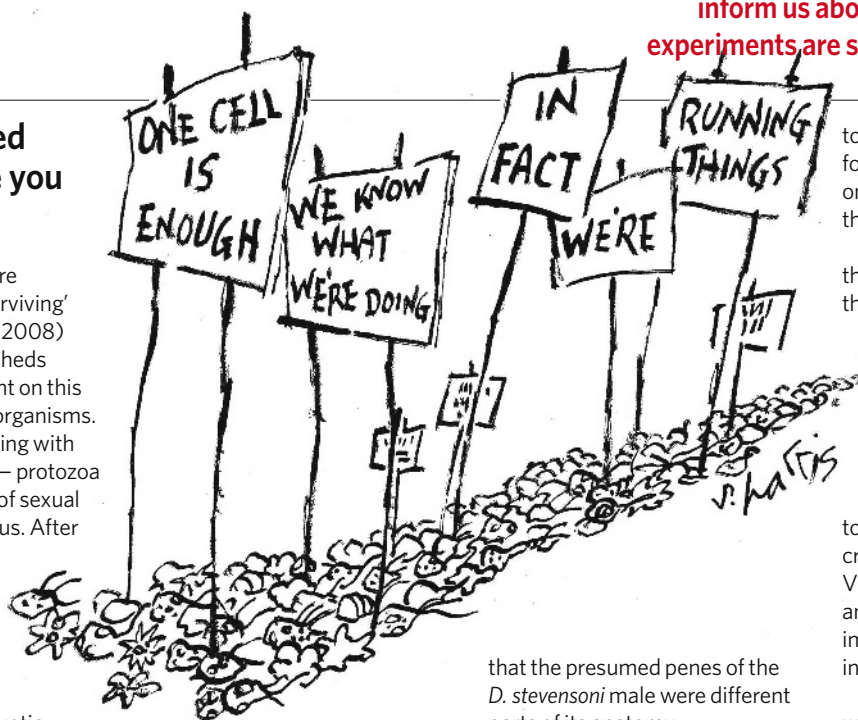
If you don't need change, maybe you don't need sex

SIR — Your News Feature 'Sex-starved and still surviving' (*Nature* **452**, 678–680; 2008) about bdelloid rotifers sheds some much-needed light on this understudied group of organisms. But for those of us working with unicellular eukaryotes — protozoa and the like — this lack of sexual habit is not so scandalous. After all, the idea that sex is widespread among eukaryotes is based on relatively large organisms.

Still, a proper understanding of eukaryotic ecology and evolution must start with unicellular eukaryotes. These drive major ecological processes and form the principal part of the eukaryotic phylogenetic tree, whereas multicellular eukaryotes are mainly figureheads. Any textbook on evolution authored by a protozoan would give large multicellular organisms and their sexual problems little more attention than a footnote.

Being eukaryotes ourselves, it is interesting for us to know why some of our relatives are sexual beings and others are not. The sexual preferences of our cousins the rotifers are probably more akin to those of protozoa than to our own. Bdelloid rotifers share many features with protozoa apart from their lack of sex life, including small body size and very large populations. The adaptational reasons for the absence of sex are not fully understood. Part of the explanation may be that the microniches these small organisms inhabit have remained unchanged for millions of years — so mechanisms for substantial genetic change are not needed.

Flemming Ekelund, Regin Rønn
Terrestrial Ecology, Biological Institute, University of Copenhagen, Øster Farimagsgade 2D, DK-1353 Copenhagen K, Denmark



Ancient asexuals: darwinulids not exposed

SIR — The paradox of sex, using bdelloid rotifers as an example of ancient asexuality, is highlighted in your News Feature 'Sex-starved and still surviving' (*Nature* **452**, 678–680; 2008). But bdelloids are not the only certain ancient asexuals — there are other contenders for the title.

The box on darwinulid ostracods does no justice to research findings on other putative ancient asexuals. Sex in darwinulids has not been conclusively demonstrated. The three males in a single species of the darwinulid genus *Vestalenula*, found among thousands of females, did not have identifiable sperm (nor did any of the investigated females), suggesting that these specimens are non-functional atavisms.

Perhaps the most important result to emerge from the discovery of these males is that a ghost male of the model darwinulid species *Darwinula stevensoni*, supposedly described in 1870, can now be dismissed, because the morphology of the copulatory appendages of the Japanese males indicates

that the presumed penes of the *D. stevensoni* male were different parts of its anatomy.

Neither does *D. stevensoni* bear any genetic signature of recombination. You correctly note that it has very low nuclear genetic variability, but fail to mention that this is expected of any organism, asexual or sexual, that has active gene-homogenizing mechanisms, such as gene conversion or efficient DNA repair.

On the basis of the available evidence, bdelloid rotifers are ancient asexuals. Likewise, the species *D. stevensoni* has now been asexual for about 25 million years. To dismiss all darwinulids as putative asexuals is as wrong as claiming ancient asexual status for the group as a whole.

Koen Martens, Isa Schon Royal
Belgian Institute of Natural Sciences, Freshwater Biology, Vautierstraat 29, 1000 Brussels, Belgium

Time to break the silence around virtual-water imports

SIR — Your News Feature 'More crop per drop' (*Nature* **452**, 273–277; 2008) mentions that China has "unconsciously" turned to 'virtual-water' imports by importing food that requires large amounts of water to produce. Officially, however, the Chinese government continues

"Understanding our auditory pathways may inform us about why only some musical experiments are successful" Laurel Trainor, page 598

to advocate self-sufficiency in food because it regards reliance on international food import as a threat to domestic security.

In spite of this strict policy, the annual virtual-water import through food trade increased sharply from 30 billion cubic metres in the 1990s to an average of 71 billion cubic metres a year between 2000 and 2004 (J. Liu *et al.* *Water Int.* **32**, 78–90; 2007).

The increase is mainly due to the import of water-intensive crops, particularly soya beans. Virtual water is politically silent and economically invisible, and importing it is current practice in China.

China is confronted with water scarcity in several parts of the country, particularly in the North China Plain and the northwestern regions. Several studies have been published on the benefits of incorporating a virtual-water strategy in regional-water management and in food-trade policies for arid regions. These all indicate that importing virtual water into the North China Plain may well be more efficient than transferring 'real' water through the South-to-North Water Transfer Project — a controversial initiative now under construction to divert water from the Yangtze River to northern China (H. Yang and A. Zehnder *Water Resour. Res.* **43**, W12301; 2007).

If a virtual-water strategy had been taken into account in a feasibility study, the decision to invest half-a-billion euros in the transfer scheme might have been different. It would benefit China's development if the political silence around virtual-water transfer were broken and if the issue received more attention in planning water resources.

Junguo Liu Swiss Federal Institute of Aquatic Science and Technology (Eawag), Überlandstrasse 133, PO Box 611, CH-8600 Dübendorf, Switzerland
H. H. G. Savenije Delft University of Technology, Department of Water Management, PO Box 5048, 2601 DA Delft, The Netherlands

COMMENTARY

Retiring retirement

The United States and Australia have done away with this archaic practice. **Peter Lawrence** says it is time to end mandatory retirement worldwide.

The evening's the best part of the day.

— *The Remains of the Day*, Kazuo Ishiguro

One of the winners of last year's Nobel prize for medicine or physiology is Oliver Smithies, an 82-year-old Englishman who emigrated to North America in 1953. Had he stayed in Europe, Smithies would have been forced to retire some 17 years ago. Instead, in the Chapel Hill School of Medicine at the University of North Carolina, he works alongside younger scientists, holds grants and exudes youthful enthusiasm. Francis Crick also emigrated to the United States at 60 and worked there into his eighties; like Barbara McClintock, Max Perutz and many others he worked right up to his last days. And these aren't simply hangers-on. While at the Salk Institute in La Jolla, California, Crick, best known for his work in molecular biology, started a "second career as a neuroscientist, publishing more than two dozen highly influential and highly cited theoretical papers pursuing the neuronal basis of dreams, memory, dendritic spines, cortical neuroanatomy and, of course, visual consciousness," according to collaborator Christof Koch at the Salk. Sydney Brenner was recently hired at the age of 80 to work at Janelia Farm Research Campus, Howard Hughes Medical Institute's new venture in Ashburn, Virginia.

These are just a few examples of the men and women who continue to make valuable contributions in their later years. In the United States, older people do all kinds of jobs, some because they enjoy working, others because they need the money. The fundamental reason, however, is that unlike in Europe, all have the right by law to be considered for work, independent of

age. Measures against age discrimination also operate in Australia and Canada. But in Europe and Japan, mandatory retirement policies continue and institutionalize discrimination.

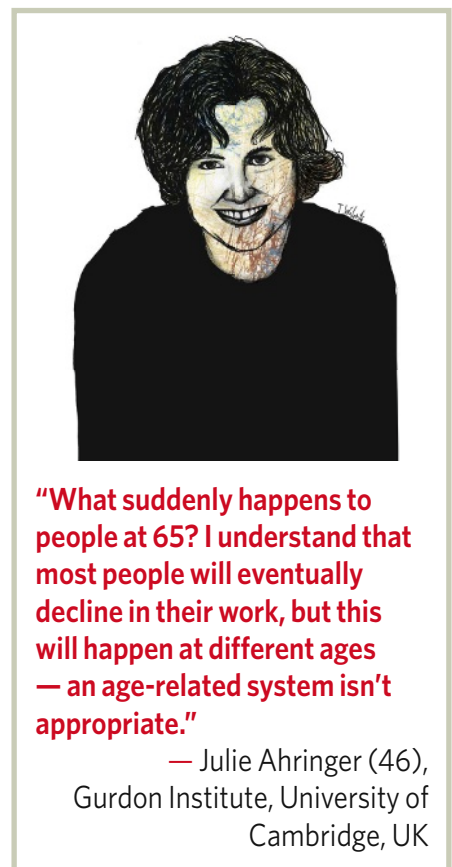
The contributions of scientists well beyond their sixth decades falsifies the case for these prejudicial policies. Indeed, German President Horst Köhler recently spoke out against mandatory retirement and described the major gains in life expectancy of the past century as "a great gift to us all". What a waste it is to shove older citizens away into retirement, he said. "We could achieve much more if we allowed the curiosity and impulsiveness of youth to be tempered by the wisdom and inner calm of the old."¹

Defunct ideas

For twenty-five years he's been lecturing and writing about things that any intelligent person already knows, and no stupid person cares to know ... which means that for twenty-five years he's been keeping somebody else out of a job.

— *Uncle Vanya*, Anton Chekhov

Prejudice is natural to the human mind: we assess people as groups and we stereotype individuals. For example, many tend to focus on the real (but small) differences between the abilities of men and women while ignoring the real (and large) differences between individuals, independent of their gender². In recent years we have become more sensitive to discrimination in some of its forms, and we now accept that, even though most women are not as strong as the average man, it would be unjust to turn down all female applicants for jobs as baggage handlers. Nevertheless, in



ILLUSTRATIONS BY T. VALENTA

"What suddenly happens to people at 65? I understand that most people will eventually decline in their work, but this will happen at different ages — an age-related system isn't appropriate."

— Julie Ahringer (46),
Gurdon Institute, University of
Cambridge, UK

"There are few middle-aged and almost no older scientists who actually work in the lab. At 80, I am one of those few. For me there has never been anything unusual about wanting to experiment at the bench. I thank my lucky stars that I was not forced to retire."

— Joseph Gall (80), Carnegie Institution
of Washington Department of Embryology,
Baltimore, Maryland



many countries, it still seems natural to deny people older than 65 the right to be considered for a post.

Many say that scientists lose their enthusiasm with age. This is certainly not always true. David Anderson, at the Howard Hughes Medical Institute, and Brenner marvel in a coda to the wonderfully original career of Seymour Benzer, at how the departed fly geneticist went on "sipping hot tea, munching Fig Newtons and asking rare but razor-sharp questions well into his mid-eighties"³. Eric Davidson, now 71, and a professor at the California Institute of Technology in Pasadena, recalls with fondness the vigour of his predecessor Edmund Wilson: "He wrote his wonderful third edition in his late sixties and it was published when he was 69 (1925). He continued to be active into the late 1930s. History shows that Columbia University, even in that benighted period, had a better policy of valuing their scientific

treasures than does modern Europe.”

So, how did the United States achieve its open-minded and liberal attitude towards older people and how well has it worked?

The American experience

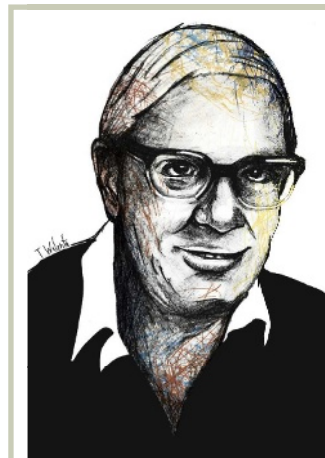
Mandatory retirement is an extravagant waste of people pie... economically, socially and spiritually, it's wrong.

— Claude Pepper, *Time Magazine*⁴

The liberal US senator Claude ‘Red’ Pepper (Democrat, Florida) (1900–1989), was known as a champion of the elderly, but famously fought against all discrimination. His 1945 bill determined equal pay for men and women. And in 1978, another of his bills abolished mandatory retirement for federal government workers and raised, from 65 to 70, the age at which non-federal employees could be forced to retire. The passing of that law was a famous occasion in the senate: to celebrate, Pepper filled Congress’ great hall with feisty septuagenarians. This was just a stepping stone; Pepper’s next bill in 1986 made all age discrimination illegal, and although academic institutions were able to delay implementation, even they complied by 1994.

Biologist John Bonner (88) remembers “The agony Princeton University and all other institutions went through when it was first announced: It will be unfair to the younger generation; there will be a bunch of doddering idiots standing in the way of progress, and so forth. Well, none of that happened.” Instead, Pepper’s initiative and others like it in Australia resulted in a scientific exodus from Europe and Japan. One example is Chris Nordin, who left a director’s post at a Medical Research Council unit in the United Kingdom in 1981; at 61 years old he was only 4 years away from mandatory retirement. Twenty-seven years later he is still working in Adelaide.

In the United States, older scientists make various contributions. Some are great role models and mentors, some augment the international reputation of their institutes,



“Unless you have seen the development of a field, all ‘facts’ can appear equally established. And with techniques: many young researchers nowadays choose a cloning method to approach their problem, letting fashion dominate the way they think. Older technologies — such as those using radioisotopes — may not occur to them. Experience is vital in making the right choice.”

— Mark Bretscher (68), MRC Laboratory of Molecular Biology, Cambridge, UK

some teach or administrate, freeing younger scientists. More importantly, they can provide a deeper perspective on scientific strategy — many young scientists have not grasped the importance of seeking out unsolved and unregarded problems.

In the same ways that academic society has been rewarded by reducing discrimination against women, increasing the presence of the elderly has also made its mark. And both practices ensure dignity and justice within the enterprise of science. As Brenner puts it: “I don’t want to retire to play golf. Science is one’s hobby and one’s work and one’s pleasure.”

Meanwhile the bad effects of compulsory retirement are multiple and insidious. Years before retirement, it turns able academics into lame ducks: they cannot take on commitments such as graduate students, and they lose their negotiating power because they cannot seek new posts. It also drives gender disparity: in 2006 in the United Kingdom, 85% of men reaching retirement age qualified for a full pension (having worked for 40 years); only 35% of women did so⁵. Although reforms set for 2010 should improve that situation for women in the UK, qualifying for a full pension remains difficult in science; training is long and many people move from country to

country. Thus, places that employ mandatory retirement have been forcing many people, particularly women, into unemployment on an inadequate income.

Managing the abolition

I know many 35-year-old scientists who should be retired and some 70-year-olds who are the best postdocs you will ever find.

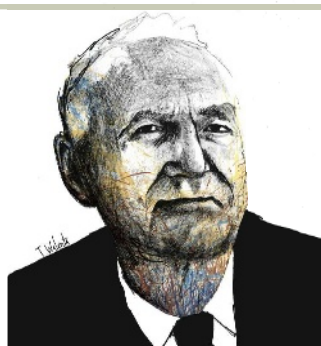
— Sydney Brenner

Age may bring wisdom, but not every ageing scientist earns their keep. This fact raises two important arguments for mandatory retirement, both with complex consequences. Firstly there is the difficulty of evaluating scientists and making them leave when their performance no longer justifies their space or costs. The contribution of people should be judged in an open-minded and tolerant way, not just by counting papers⁶ or by undergraduate ratings on teaching, but by evaluating the entire contribution to the institution. If following this assessment, old (or young) are fairly judged to be incompetent, inactive, disruptive or wasteful, they should leave, or have their lab and office space reduced to an appropriate level — to do otherwise is to damage the institution and deny others opportunities. Research grants won or lost can serve as arbiter; however, this is an external and subjective judgement that does not take into account teaching, mentoring, administration or other contributions. The problem of selection has to be faced: the traditional alternative, to throw everyone out at a certain age, regardless of their wishes or their usefulness, simply because it avoids difficulties of management, is indefensible.

The second difficulty is financial. It is said that older scientists can be so expensive that insufficient funds remain to hire new personnel. Biochemist Peter Hartmann (67), at the University of Western Australia in Perth is at

“I had four papers last year and have more in the pipeline. I am now Chairman of the South Australian Working Party on Osteoporosis and Fracture Prevention. Every year, my accreditation to the Hospital has been renewed but I hope I will give up before I am asked to do so. There is still a lot left to do.”

— Chris Nordin (88), Royal Adelaide Hospital, Australia



least one counter example: "The direct cash contribution that I make to the university from my research grant is about 50% greater than my salary. In addition I pay 80% of a lecturer's salary to cover my teaching and administration obligations. Thus if I were to retire there would be a negative financial benefit to the University."

The solution

To address the general situation, we must be radical and flexible. If scientists take on fewer duties, it makes sense to reduce their rewards. Furthermore, many scientists have built up pensions; in my opinion these pensions should be started on time and the salary reduced partially or completely. If the scientist is of particular use to an institute, it may wish to reward his or her contributions by some supplement to the pension. This will vary depending on how useful the person is.

Remember, we are not discussing simply the right of an employee to go on working, rather the right of every individual to be considered for work on merit and negotiate for remuneration. A portion of anyone's salary is an investment in the future, to retain valued teachers and researchers, and therefore the justification for that portion will wane in older scientists until, sometimes, it may reach the point when the main beneficiaries of the opportunity to work may be the scientists themselves. The logical result of these arguments is to restructure salaries so that they can fall as well as rise.

Let's look at a real case: the University of Cambridge is currently tussling with these problems because, thanks to pressure from the European Union, the law is inching towards a change in the United Kingdom. As of now the normal retirement age for teaching staff is 67

or 68 and, when an extension is requested, the university may currently offer up to three more years. More than 50% of staff are applying for an extension, presumably, among other things, reflecting a high level of job satisfaction.

The university is considering each case individually and awarding variable contracts, depending on the duties agreed. In this way it is exploring its options, rather as universities in the United States have done. In all cases the primary determinant is the opinion of the relevant department; if the head of that institute wants to keep the person, the university tries to support them — if it can find the money. This makes good sense as a centralized department of human resources could not make these individual judgements.

This process has only been running for 18 months, so the extended contracts have not yet come to an end. But when they do, in a year and a half, the university has indicated that people will be able to apply to have them extended further. Experience in the United States suggests that fewer will want to go on much beyond their mid-seventies⁷.

One way to take more account of individual variation is to free up the whole process further by weakening academic tenure, which up to now has been sacrosanct. We have to be realistic and admit that offering secure tenure with a salary for 30–40 years regardless of whether there is any useful contribution to the employing institute is wasteful and unfair. Indefinite tenure has already been abandoned by many government research institutions.

John Bonner speaks for many when he insists that every institution ought to "evaluate each elderly person and provide them with the means of continuing their work if they remain productive and are driven in that direction".



"What is the point of starting something new and exciting in your last few years, if you know you won't be able to continue the work? One would spend one's last few years dotting the 'i's and crossing the 't's."

— Cynthia Kenyon (54),
University of California, San Francisco

It is high time to do elsewhere as the Americans have already done: to allow the pursuit of happiness to those proficient older citizens who wish to seek or hold employment.

Peter A. Lawrence (66) is in the Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, and the MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK.
e-mail: pal@mrc-lmb.cam.ac.uk

1. Köhler, H. Christmas address, www.bundespraesident.de/rede-2.642777/Ein-neues-Miteinander-von-Jung.htm (2007).
2. Kovas, Y., Haworth, C. M., Dale, P. S. & Plomin, R. *Monogr. Soc. Res. Child Dev.* **72**, 1–44 (2007).
3. Anderson, D. & Brenner, S. *Nature* **451**, 139 (2008).
4. www.time.com/time/magazine/article/0,9171,915567,00.html
5. www.dwp.gov.uk/pensionsreform/pdfs/GenderImpactAssessment-5-Dec2007.pdf
6. Lawrence, P. A. *Curr. Biol.* **17**, R583–R585 (2007).
7. Ashenfelter, O. & Card, D. *Did The Elimination Of Mandatory Retirement Affect Faculty Retirement?* American Economic Review, National Bureau of Economic Research **92**, 957–980 (2002).

In the course of researching this article, the author communicated with more than forty scientists, compiling their thoughts and musings. For more communications and to respond with your own, visit <http://tinyurl.com/4mkdvu>.

"When I was young, say in my thirties, research grants for young scientists were very poor in Japan and we complained. Then our government decided to support younger scientists, for which I was very pleased. But I was almost immediately categorized to be 'too old' for new grants. So I jumped from being too young to being too old, and was never the right age. It is an injustice and a waste if active scientists are thrown out of work simply and only because they reach a certain birthday. Some are sharp at over 90, some are dull even when physically young."

— Motonori Hoshi (67), The Open University of Japan, Chiba City



BOOKS & ARTS



M. EVANS/AP PHOTO

Vivisectionists strike back

The complex issue of animal experimentation should not be dumbed down in the face of violent opposition, argues **Andrew Read**, reviewing a new polemic on the US experience.

The Animal Research War

by P. Michael Conn and James V. Parker

Palgrave Macmillan: 2008. 224 pp. \$34.95, £22.99

Animal researchers in the United States are the targets of terrorists. If you need persuading that this is happening and that this is bad, read this book. It is a primer on the extremists and organizations involved and their tactics — lobbying in schools, novel legal manoeuvres, threatening families, posting razor blades, vandalism, arson and bombing. If you experiment with animals, this book should persuade you to open your post with a letter-opener instead of your finger. There are people out there who really hate you. Authors Michael Conn, associate director of the Oregon National Primate Research Center, and James Parker, the centre's former press officer, argue that extremists are waging a war on animal research.

Talk of war can motivate a nation, but it does not always generate good policy. Conn and Parker say they must use military rhetoric to attract the attention of the public to these attacks and their effect on human and animal well-being. To deliver their wake-up call, the authors mirror the strident language of animal-rights organizations, such as People for the Ethical Treatment of Animals (PETA). I have some sympathy for the authors' approach. The enormous membership of such organizations

suggests that the best way to get a message across in this dumbed-down world is to skip the subtleties, assert that grey is black, and hurl sound bites, anecdotes and context-free quotes at the public. *The Animal Research War* does this important job well enough.

But tossing certainties over the barricades will go only so far. To agree rules we must discuss the details, and the ethics are difficult. Animal experimentation undoubtedly contributes to relieving human and animal suffering, so we are morally obliged to do it. But we are also morally obliged to do it humanely. Achieving both these ambitions with public support requires a grown-up debate.

No one knows what level of governmental oversight best reconciles our moral obligations to the sick with our moral obligations to animals. The US system is less demanding than that in the United Kingdom, for example, so which is right, if either? At least one approach must be morally deficient and should be abandoned.

Another difficult issue is how we should deal with phylogeny. In the United States, rats, mice and birds are not given the same legal cover as primates, cats and dogs. Under the UK system, one species of octopus is covered, but you can do what you like to the others. Is there a sound basis for replacing mice with zebrafish in toxicology testing? Should there be welfare protection for crustaceans? Or no protection for chickens?

In most countries, permission to experiment

on animals involves an ethics committee that performs cost-benefit analyses. Yet the connection between a particular experiment and human well-being is usually tenuous. Science is one of mankind's most wasteful enterprises, and it has to be: as Conn and Parker put it, scientists are explorers, not clairvoyants. Research that seems to be useless, and even research that actually was useless is a crucial part of the long-term success of biomedicine. Many animals are used to explore what turn out to be blind alleys. That is how we learn.

Is it even possible to determine the cost of not experimenting on animals? One of the reasons that malaria continues to be such a humanitarian crisis is because, compared with primates, mice can mount much better protective immune responses against malaria. Ideally, high-throughput screening of candidate vaccines should therefore be done in primates. How much human suffering is a consequence of the systematic failure to use more primates in malaria vaccine development?

These issues are at the heart of animal experimentation, and the scientific community is silent about them. Do we avoid open discussion about national regulations because these rules shield us from society? One of the worst consequences of animal-rights extremism is the suffocation of rational discussion. If complex arguments are not aired, then the difficult decisions are left to politicians who must juggle vested interests.

The scientific community, one group with such vested interests, is now being forced by the political activism of the antivivisectionists to clamber out of its bunker and engage, at least with the easy questions, such as why animal experimentation cannot be banned outright. *The Animal Research War* is an important US step in the hitherto largely European process of making that simple case. But in the online era, the role of books is to provide accessible analysis of the details behind the sound bites, and this book should have gone further. For instance, Conn and Parker's research centre was infiltrated by an undercover animal-rights activist, and pictures of its apparently miserable primates are posted on PETA's website. Conn and Parker counter that the pictures were either doctored or taken deliberately just before cages were cleaned, or show behaviour seen naturally in the wild. So why not give visual proof? Why aren't we letting the cameras in?

Animal experimentation is a complex issue. Pretending otherwise, as Conn and Parker do, smacks of a cover-up. It is farcical to maintain that biomedical scientists are always angels, particularly if our best argument is that the US government has given us the all-clear to experiment. Let's be honest about past abuses and new cases that come to light, and move forward with case studies that show how animal research is run better these days. There may be mileage in animal-centred historical accounts being hurled at us by animal-rights campaigners, which Conn and Parker simply throw back with the rhetoric reversed. Is the correct lesson of thalidomide that animal testing is useless, or that insufficient animal testing is dangerous? Would the germ theory of disease still be conjecture without animal experimentation? Where would the polio vaccine be if Sabin and Salk had modern *in vitro* and *in silico* tools and twenty-first-century animal regulations? The undoubted cases of excessive

and insufficient animal use, the dead ends and blind alleys, and the basic biology from whence it all came may be rich in honest detail, but such narratives and conjectures would make fascinating and persuasive reading.

The full story would also address the fundamental misunderstanding at the heart of the debate. Antivivisectionists maintain that researchers actively choose to experiment on animals, despite personal risk and bureaucratic hassles, to pursue science that is irrelevant to human or animal well-being. They also believe that our community conspires to keep the pointlessness of our profession from the public. Actually, our biggest problem is that we are poor at projecting the complex and messy process of science, and the humanity of scientists. ■

Andrew Read is professor of biology and entomology in the Center for Infectious Disease Dynamics, Pennsylvania State University, Pennsylvania 16802, USA.

Moralist, meet scientist

Experiments in Ethics

by Kwame Anthony Appiah

Harvard University Press: 2008. 288 pp.
\$22.95, £14.95

Picture a crowded room. In one corner, people are arguing about abortion; in another, about euthanasia. Around the coffee table, sitters dispute the obligations of the rich to the poor. By the sofa, folks are debating the criteria for a just war and the proper relations between men and women, and another group queries the use of primates in medical experiments. Nearby, a huddle of ethicists disputes politely which moral theory to accept — utilitarianism, kantianism, virtue ethics, contractarianism or something else. At the back, people are shouting something about whether suicide bombers are heroes or villains, and plates are flying. The individuals seem to have only one thing in common: each is convinced he is right.

A new guest arrives. Finding the front door unlocked, Science — not in the habit of knocking — has barged in on the pandemonic party. Will Science resolve the disputes and settle who is right? Will Science make us all look silly, showing we are squabbling over words that have no meaning? Or will Science remain aloof, like a nerdy stranger, observing the goings-on yet unable to address the rambunctious crowd?

In the past few decades, scientific interest in moral behaviour has surged. Psychologists,

neuroscientists, evolutionary theorists and behavioural economists have begun to turn their experimental methods to understanding the ways we arrive at moral judgements. Scientists of human nature have called into question commonplaces about character and offered subversive explanations for various moral intuitions.

Philosopher Kwame Anthony Appiah explores the relation between empirical research into moral behaviour and moral philosophy, a discipline that questions what we ought to do and what there is reason to value. In *Experiments in Ethics*, he reviews a sample of the most intriguing experiments through which scientists have sought the mechanics of our moral minds.

Questionnaires have revealed that people's responses to moral dilemmas sometimes

depend on how a problem is framed. This calls into question how much weight we can put on moral intuitions in cases in which superficial rewording makes us reverse our verdicts. Hypothetical problems in decision-making have demonstrated that general biases in human thinking arise when cognitive heuristics are applied outside their proper domains. Daniel Kahneman received the 2002 Nobel Prize in Economics for developing prospect theory, describing how people's behaviours deviate from the prescriptions of classical decision theory. Similar biases might occur in our moral thinking. In a famous study on moral judgement that used functional magnetic resonance imaging, Joshua Greene reported that utilitarian and non-utilitarian responses were associated with different neural signatures in moral 'up-close-and-personal' dilemmas.

In one philosophical thought experiment, a runaway trolley threatens to run over and kill five people. You can flick a switch that will deflect the trolley down a different track where it will kill one person. Most people say they would flick the switch — a response that is in accordance with utilitarianism, an ethical theory according to which the right action is the one that has the best net consequences.

Now consider a variation: again a trolley rolls along a track where it will kill five people. This time you are standing on a footbridge overlooking the track, and the only way you can save the five is by pushing a fat man standing beside you down on the track; if you do, the fat man will die, but his body mass will stop the trolley and five people will be saved. In this more personal version, most people say they would not push the fat man even though



A passer-by is more likely to offer help if they have recently experienced good fortune.

fewer people would die if they did.

Greene found that the brains of subjects who gave the non-utilitarian response — not to push the fat man — in this case showed increased activation in regions associated with emotion, whereas the brains of the rarer utilitarian responders were activated in areas associated with working memory and reasoning. This finding has been interpreted as support for utilitarianism, the theory supposedly reflective of more rational cognitive processes. Others dispute this. In a recent study by Michael Koenig and colleagues, patients with damage in brain areas involved in the normal generation of social emotions were more likely to select a utilitarian response. One can imagine ways of spinning this result that would be less flattering to the utilitarian.

Other investigations focus on behaviour. Dropping your papers outside a phone booth in a shopping mall, you are much more likely to get help from passers-by if they have just found money in the phone's coin-return slot, according to one study. Placing a coin in the slot increased the proportion of people who stopped to help, from almost nobody to almost everybody. You are also much more likely to get change for a note from a passer-by if standing near a fragrant bakery shop than outside a "neutral-smelling dry-goods store". Many studies have confirmed that morally irrelevant features have a strong effect on whether people offer help. Some have even concluded that character is a myth: it is the situation, rather than any stable personality trait, that determines whether one does the right thing. If true, this claim poses a big challenge to aristotelian ethics, which is built on the idea that we ought to cultivate a virtuous character. However, causes can be multifactorial: the fact that situational features often have a big influence does not imply that character never has any influence. It is probably a question of degree.

Psychologists have also sought to identify the basic 'taste buds' of our moral sensibilities. Jonathan Haidt of the University of Virginia, Charlottesville, proposes we have a limited number of distinct moral modules, which get activated in response to different kinds of moral problems. Haidt's taxonomy distinguishes responses related to harm, fairness and reciprocity, hierarchy and respect, purity and pollution, in-group and out-group boundaries, and awe and elevation. If we do have all these sensibilities, one might suspect that moral theories that recognize or privilege only one or two of them are unduly restricted.

Experiments in Ethics is erudite, concise

and engagingly written. Appiah assesses that experimental science is relevant to the enterprise of normative ethics, and that the relation between the two, although complex, need not be antagonistic. Returning to my analogy, Appiah recommends that our lab-coated gatecrasher enter into a dialogue with the party-goers, and that they should welcome Science into their midst. The nine

Muses of music and the arts should also be invited. Seeing the central issue of ethics as eudemonia, the good life, Appiah believes that such a plenary gathering will best enable both its study and its achievement. ■

Nick Bostrom is director of the Future of Humanity Institute, Faculty of Philosophy and the James Martin 21st Century School, University of Oxford, Oxford OX1 4JJ, UK.

Music grown from garden weeds

Umbel Ballits: Dylan Martorell
Craft Victoria, Melbourne
Until 28 June

Botany, traditionally the preserve of water-colourists with a meticulous eye for detail, continues to inspire contemporary artists. Graphic designer and musician Dylan Martorell uses plant morphology as a starting point to create sound sculptures. His latest work is now on show in Melbourne, Australia.

Martorell begins by drawing plant-growth algorithms in pencil or ink on paper. Based on patterns he discerns in botanical structures, he uses a visual notation that enables the diagrams to function as musical scores. For this exhibition, he sourced plant structures from an Australian botanical manual by David Whibley and Trevor Christensen called *Garden Weeds — Identification and Control* (Botanic Gardens of Adelaide, 1991).

Moving beyond his precise drawings, Martorell also constructs less ordered three-dimensional sculptures. Evoking the original botanical structures, he incorporates found objects, musical instruments and living plants. He also digitally converts data from his drawings into music that slides from one frequency to another. Selections of the resulting 3.5-minute sound tracks are played simultaneously in Craft Victoria's gallery, generating complex, overlapping sound patterns that create a fourth, aural dimension to his multimedia

interpretation of the plant world. A small ensemble of musicians, including Martorell, occasionally rehearses and performs live music in the gallery, using the sculptures as instruments to generate sound.

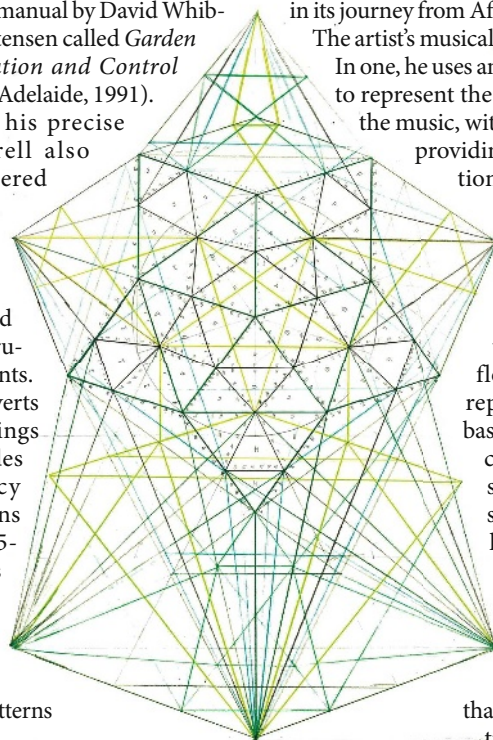
"Currently I am creating hand-written scores based on the floral structures in various flowering plants such as racemes [unbranched flowers] and umbels [umbrella-like flowers]," says Martorell. Although the title of his exhibition — *Umbel Ballits* — mimics linnaean binomial classification, that was not his main intention. "Ballits is old Afro-American vernacular for ballads," he explains. "The sound pieces in my exhibition are based on pentatonic tunings, which were influenced by a variety of musical traditions, including the development of blues music in its journey from Africa to America."

The artist's musical scores take two forms.

In one, he uses an upright trunk or stem to represent the base note or drone of the music, with branches or clusters providing extra musical notation. Scales and frequencies are represented on the horizontal axis and time is represented vertically.

His second score pattern is shaped like a flower, with each petal representing a different base note, either played in clockwise succession or simultaneously. Using structures derived from local flora as his building blocks, Martorell draws, sculpts and composes multilayered and location-specific artworks that challenge our perception of botanical art. ■

Colin Martin is a writer based in London.



Dylan Martorell creates sounds using his algorithms of plant growth as musical scores.

D. MARTORELL: WWW.HIDDENARCHIVE.COM

Reimagining the Royal Institution

The Royal Institution of Great Britain in London reopens this week after two years of renovations costing £22 million (US\$ 44 million). The eighteenth-century building now flaunts its heritage, spotlighting the scientists who worked there and discovered 10 chemical elements and claimed 14 Nobel prizes (see page 568). Architect **Terry Farrell** explains how he rethought the institution's ethos.

You are a leader in ecology-based design and city planning. How does the natural world inspire your work?

I became interested in nature as a child and was a pupil of Ian McHarg, a pioneering naturalist and ecologist who taught landscape design at the University of Pennsylvania in Philadelphia. I have a great passion for communicating ideas, and enjoy books that explain complicated things, such as town planning or the sciences, by authors including Bill Bryson, Richard Dawkins and Matt Ridley.

Science informs my work — I am design champion for the Thames Gateway, the urban regeneration project stretching from east London to the Thames Estuary, and it is my ambition for this to be Britain's first eco-region. We have also designed many aquariums, including The Deep in Hull, UK, the Seattle Aquarium in Washington state and Biota! in London.

Famous scientists including Michael Faraday, Humphrey Davy and Lawrence Bragg worked at the Royal Institution. How did its history influence the renovations?

The project was led by the energy, drive and ambition of the institution's director, Susan Greenfield. We were closely at her side helping to rethink the building and its mission statement, and in doing so have managed to update and redefine the great history and great brand of the Royal Institution.

The historic fabric of the building was tired and run-down, its foundations were failing, the classic façade was in urgent need of repair and the services were in disarray. The richness of its history, such as the Nobel prizes won, the chemical elements discovered within it and the history of its growth, change and research projects were not apparent.

A careful balance was needed between preserving the building's history while adding new elements to give it freshness and vitality. It was not just a question of refurbishment, but of creating a new feeling and energy.

Over the centuries the building grew in an ad hoc way. How did you bring cohesion?

A lot was achieved by opening up and connecting the spaces. There was no mental



A. WILLIAMS

map of the building before, owing to its confused circulation system. The grand stair went only from the ground to the first floor, and two other staircases were tucked away at either end of the building — one had been nicknamed 'the sad staircase' because it was so neglected. A major part of the work was to implement a really good circulation system running primarily through the centre of the building, giving it a heart, a focus and a point of orientation. Vertical movement has been rationalized by introducing a new central lift within a glass atrium, spanning the whole height of the building.

The mission of the Royal Institution was realized by extending the exhibition areas and using the whole fabric of the building to show the collection of paintings and drawings. Displaying objects and telling the stories behind the artefacts led to a complete rethink of the exhibition area in the basement. The Faraday Theatre has been given a cooling system for hot weather and the auditorium seating rearranged because people are larger nowadays.

How did you rethink the institution's purpose, given its many roles from research to outreach?

One of the biggest changes in society is the democratization of learning and the need for accessibility to institutions by everyone: male or female, rich or poor, young or old. This

modernity was achieved through a new kind of architecture, the transparent glass atrium, a glass lift, and through lighting, materials and freshness. And all this is done in a way that keeps the best of the old.

Extending the opening hours of the building will make visits longer, creating the possibility of meeting friends for coffee, dinner or drinks. It is a place for anyone to come and enjoy, not just for 'stuffy academic' types.

The children's laboratories and the televised annual Christmas lectures were a long-time commitment to the Royal Institution's outreach. A Young Scientists' Centre has been designed in the basement exhibition space, where children can perform experiments on heating, waste, lighting, power, gas, water, air and information technology.

The UK Science Media Centre is housed here and, through traditional and modern communication methods, will be able to respond to and be proactive on matters of scientific interest, connecting the institution's interests to the wider world.

Through all these interventions, we have reconfigured the Royal Institution not as a museum, but as a living, working, lively and engaging establishment.

Interview by **Matt Brown**, editor of <http://network.nature.com>, Nature's social networking site for scientists.

Seeing the smaller picture

Advances in imaging techniques are transforming microbiology into a science that is rich in visual imagery, harking back to biology's pre-darwinian origins, explains **Martin Kemp**.

Most sciences go through visual and non-visual phases. There are times when attempts to visualize and represent nature stand at the cutting edge. At others, measurement, statistics and algebra hold sway. The two modes are cyclical. Visual phases are often driven by the advent of technologies that allow us to travel optically into new realms, minute or vast.

Until Charles Darwin (1809–1882), biology in its macro form was the most visually rich of sciences. But Darwin's visual austerity was extreme. No pictures are included, for example, of the domestic pigeons that played a starring role in chapter 1 of *On the Origin of Species* (1859). This omission must be seen in part as a reaction against the elite picture books that dominated natural history at the time. Darwin's contemporary, Gregor Mendel, with his background in physics and statistics, used numbers in his work on inherited traits in plants.

Today's biology is again embracing the visual. Writing in his blog *Small Things Considered* (<http://tinyurl.com/3w3gx2>) for the American Society for Microbiology (ASM), Moselio Schaechter, a former ASM president, has hailed the advent of a brave new visual world in 'small biology'. "Not so long ago, it would have seemed implausible that biology would return to its origins as a visual science," he notes. "Some of the most fundamental work done now once again involves seeing shapes and forms... Nowadays, mind-blowing insights come from seeing with your own eyes."

The illustration (pictured) accompanying Schaechter's blog is one of the many brilliant drawings of the shells of foraminifera made by Christian Gottfried Ehrenberg (1795–1876), a number of which were published in his *Mikrogeologie* of 1854. A close colleague of Alexander von Humboldt, Ehrenberg pioneered the study of microscopic fossils. His drawings exhibited the minute developmental geometries that inspired the illustrations of Ernst Haeckel and D'Arcy Wentworth Thompson half a century later. But the immediate future of biological science did not lie with such picturing.

The ASM, which dates back to 1899, proclaims Louis Pasteur, Robert Koch, Alexander Fleming, Antonie van Leeuwenhoek, Joseph Jackson Lister, William Jenner and Jonas Salk as its progenitors. Of these heroes, only the great microscopist van Leeuwenhoek (1632–1723) used pictures as a major tool.

As Schaechter combatively puts it, "Many of

the founding fathers of modern biology (read molecular biology) were physicists who were used to abstractions and numbers and didn't have the urge to look at anything. Others who should have known better were probably cowed or seduced by these amazing people."

A new world of structural biology has been opened up by advances in imaging techniques. As Schaechter says, we can now "see how an enzyme works or how macromolecules interact with molecules large and small... single-molecule imaging methods allow us to visualize the tiny movements made and the forces generated by proteins or ribosomes. One can now 'see' in real time polymerases polymerizing and ribosomes translating."

The ability to measure individual molecules in living cells promises new and vivid pictures of signalling within living cells. Experiments and theoretical modelling can now re-engage in a fresh way. Because microscopy and fluorescence labelling now permits the identification, tracking and visualization of the interaction of single proteins or RNA molecules, it is possible to match theoretical models of protein folding or molecular dynamics against what can be seen.

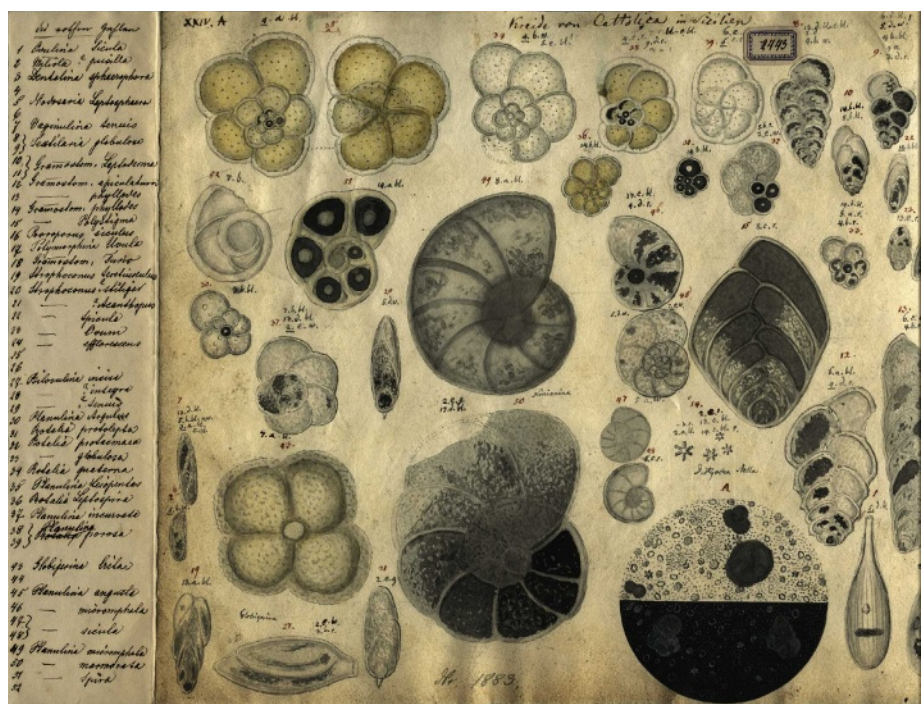
In microbiology, as in other sciences, there are competing interests at work between groups

who have commitments to particular methods and tools. Nevertheless, important frontiers in biology and biophysics are being furthered by picturing; data obtained at previously unseeable scales can be translated into vivid visual models that serve to investigate, explicate and illustrate forms and mechanisms. The dimensions are much smaller than those Ehrenberg imagined, but he would have understood the priority accorded to visual tools for understanding function through dynamic structures.

Given this visual turn, it is disappointing to see that the main ASM web pages and those for their General Meeting in Boston, Massachusetts, this week, are visual deserts. Darwinian austerity rules there. Only by following links into *Microbe Magazine* and *Microworld* does the visual flourish in the familiar form of the vividly coloured wonders of computer graphics and a series of lively videos.

Among the visual and non-visual cultures in biology, Schaechter has no doubt where the future lies: "The Age of Imaging is just beginning." He does, however, admit that "it's hard to predict where it will lead, as the limits seem to constantly recede".

Martin Kemp is research professor in the history of art at the University of Oxford, OX1 1PT, UK.



MUS. FÜR NATURKUNDE, BERLIN HUMBOLDT UNIV.

Hidden treasures: Bologna's Poggi Palace

The exquisite obstetrics models on show in an Italian palace hint at its former multidisciplinary glories.

Alison Abbott investigates the museum that was once home to a revolutionary institute of sciences.

Anatomy and art have depended on each other since the Renaissance. The bond took a new form in the eighteenth century with the development of anatomical models. Designed for teaching, as an alternative to inconvenient and infrequent cadaver dissection, these models represented the living, not the dead. They celebrated the beauty of human anatomy from the inside.

The models first emerged from studios in Bologna, Italy. Dignified, skinless men stood proud like three-dimensional versions of the Renaissance drawings of Andreas Vesalius. Models of voluptuous young women were draped provocatively on soft folds of wax-stiffened silk, the better to expose their innards to generations of medical students. Later models had more specialized functions, such as the fine obstetrics tableaux of modeller Anna Morandi-Manzolini (1716–1774), which showed pathological states for the first time.

Bologna, 290 kilometres north of Rome, is home to Europe's oldest university. For one magnificent century, it was also home to the Institute of Sciences, founded in 1711 by the army general and scholar Luigi Ferdinando Marsili, who regretted the failure of his university to engage in experimental research. His solution was to found the institute, which brought together all of the burgeoning sciences that the region offered — from anatomy and geology to astronomy and the study of electricity — in the frescoed sixteenth-century Poggi Palace on the outskirts of town. Marsili donated his own extensive scientific collection, gathered during his military travels, to the new institute. He generously equipped six laboratories and added a formidable astronomical tower to the palace. Interdisciplinary approaches would yield the greatest scientific advances, he reasoned.

His endeavour gained the crucial support of the Bolognese archbishop who, in 1740, became Pope Benedict XIV. The enlightened Benedict XIV believed that scientific discoveries would illustrate rather than diminish God's greatness. He exempted the institute's library from adherence to the Catholic Church's notorious list of prohibited books. In 1742, he commissioned and paid for artist Ercole Lelli to produce a series of life-sized wax models for the institute, to demonstrate different layers of muscle on the skeleton. A new era of scientific modelling was born.

Until the eighteenth century, obstetrics was



MUS. PALAZZO POGGI, BOLOGNA

almost exclusively the domain of midwives. Then, new instrumentation — forceps to pull out a baby without damaging its head, hooks to turn it in the womb, and macabre instruments to extract dead babies — brought childbirth into the domain of men. Naturally, the instruments demanded a better knowledge of the topology and anatomy of the stages of birth.

This prompted the obstetrician Giovanni Antonio Galli to build up a private teaching collection of models, with the help of Lelli's students. He commissioned Morandi-Manzolini to make her precise wax tableaux. Many other items were made of clay, a cheaper material. One major series (pictured) allowed medical students and midwives to view or, when blindfolded, to feel the different phases of pregnancy through to delivery. Some models demonstrated difficult and potentially deadly fetal positions, such as breach or tangled in the umbilical cord.

In 1757, Benedict XIV bought the entire collection of 170 models and donated it to the institute's school of surgery, appointing Galli to teach obstetrics. Midwives had to enter the school through a back door. When Galli died in 1782, Luigi Galvani — better known for discovering the electricity generated in animal muscles

and nerves, and already president of the institute — took over his position. Galvani also proposed the correct notion that uterine contractions during labour have a strong nervous component.

Galvani was dismissed from all of his posts in 1797 after he refused to swear an oath of allegiance to the occupying Napoleonic troops. He died in poverty the next year. During the university reforms that Napoleon enforced in the early nineteenth century, the holdings and equipment of the institute were dispersed.

In the past decade, many of the original collections and instruments, including the anatomy and obstetrics models, have been tracked down and restored to their original rooms in the Poggi Palace, which opened again in 2000. Among the many exhibits, visitors will find the lab where Galvani did his famous experiments on frogs and Ulisse Aldrovandi's sixteenth-century natural history collection, which went against the trend of the randomly organized *Wunderkammer*, or cabinets of curiosity. ■

Alison Abbott is Nature's Senior European Correspondent.

See www.museopalazzopoggi.unibo.it for further details. For more Hidden treasures see www.nature.com/nature/focus/hiddentreasures

ESSAY

The neural roots of music

Laurel Trainor explains how the emotional power of music depends on the structure of the ear, and on our basic encoding of information.

In search of an original voice, the dominant composers of the mid-twentieth century — Arnold Schoenberg, Pierre Boulez and their disciples — rejected the tonal and rhythmic forms of the past. They adhered to rigorous compositional techniques such as the serial tone-row method — in which all notes of the chromatic scale occur equally often in a repeating row — banishing tonality. Some powerful compositions were written in the serial style, but few are played regularly today. Asked in 1999 why this might be, Boulez responded: “Well, perhaps we did not take sufficiently into account the way music is perceived by the listener.”

Understanding the structure and development of our auditory pathways, and how experience modifies them, may inform us about why some musical experiments are successful, and others are not. Why, for instance, the melodies of Gustav Mahler, the driving rhythms of Igor Stravinsky or the dissonances of John Adams make these modern composers popular today, whereas the music of some others, such as Luigi Dallapiccola and Luigi Nono, is rarely heard.

Music is conceived by our brains, played through our bodies, perceived through our sensory organs and then interpreted by our brains. Thus it is subject both to general constraints of our neural system and to specific constraints of our auditory processing capacities.

During childhood, each of the billions of neurons in the human auditory system forms thousands of connections to other neurons, creating neural networks. Genes control the characteristics of neural circuits, developmental waves of neuronal and synaptic proliferation, and the later pruning of neural connections to form efficient circuits for processing sound. Experience also profoundly affects the neural connections formed. Studies show that rats raised in environments containing only white noise with no pitch or rhythm are unable to recognize everyday sounds, and are greatly impaired even in their ability to discriminate different pitches¹.

Certain sounds elicit specific, powerful emotions in people, presumably a testament to the evolutionary heritage of our auditory systems. Low, loud, dissonant sounds evoke fear; rapid, higher, consonant sounds evoke friendliness or joy. Mothers around the world talk and sing to infants using a cooing tone of voice and higher pitch than when interacting with adults. Infants prefer these higher-pitched vocalizations and

mothers sing in different styles to help prelinguistic infants regulate their emotional state. Across cultures, songs sung while playing with babies are fast, high and contain exaggerated rhythmic accents; lullabies are lower, slower and softer.

Talking to people of all ages, we use falling pitches to express comfort; relatively flat, high pitches to express fear; and large bell-shaped pitch contours to express joy and surprise. Hearing music with an unfamiliar structure, listeners base their emotional reactions largely on such sound features.

Music is built on general, universal features of human sound processing that have deep evolutionary roots. It also incorporates rhythmic, melodic and harmonic structure. Musical structures and styles vary enormously across cultures, and change as continually as languages, yet our biology constrains the possibilities.

Rhythm is a dancer

Musical rhythm may have its origins in the motor rhythms controlling locomotion, breathing and heart rate. Babies receive correlated sound and movement input as parents rock them while singing. This and other early experience encourages movement and auditory representations to wire together in the brain.

Although music makes us want to move to the beat, movement evolved first and there are multi-sensory connections in the brain between motor and auditory areas. Therefore it should also be the case that how we move affects how we interpret rhythms. We have shown this with a repeating 6-beat rhythm pattern with no accents that can be perceived as two groups of three beats (as in a waltz) or as three groups of two beats (as in a march). Adults and infants who bounce up and down on every second beat report hearing — or in the case of the infants prefer — a march. Those bouncing on every third beat hear a waltz².

The evolutionarily ancient vestibular system for balance plays a crucial role in the interaction between movement and the perception of musical rhythm, indicating that music and dance could have evolved together. Stimulating someone's vestibular nerve alternately on the left and right sides gives a sensation that the head is moving from side to side. Such stimulation alone, on either every second beat or on every third beat of an unaccented 6-beat rhythm pattern, biases judgement of whether

the music is heard as a march or a waltz³.

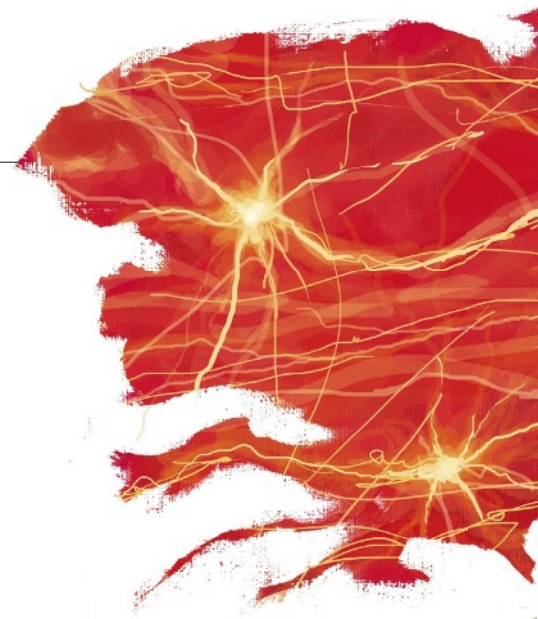
Western music, from pop to classical, tends to use simple rhythmic structures. Folk music from many other traditions uses complex structures, for example pitting groups of 7 and 11 beats against each other. Young infants can perceive complex rhythmic structure, but they lose this ability before they are a year old if not exposed to such rhythms.

Our capacity for processing rhythmic complexity is likely to underlie the comparative success of rhythmic experimentation over pitch experimentation in the twentieth century. It might explain why rhythmic structures from many traditions around the world have been successfully incorporated into music, why audiences embrace the jazz rhythms of composers such as George Gershwin, and why people find the rhythmic richness of composers such as Stravinsky, Béla Bartók, Jennifer Higdon and film composer Danny Elfman so compelling.

Pitch invasion

The use of pitch in music reflects the constraints of auditory mechanisms for identifying objects and separating sound sources. These constraints are evident in the near-universal use of consonance and dissonance as an organizing principle; in the use of scales comprising a small set of pitch categories that repeat at octave intervals, reducing the amount the listener has to remember; and in the use of at least two different-sized intervals in scales, allowing the emergence of different relationships between tone pairs and tonal functions such as ‘tonic’ and ‘dominant’ in Western diatonic scales.

Sensory dissonance arises from how the basilar membrane vibrates in the cochlea of the inner ear, and from the firing patterns of auditory nerve fibres that this movement activates. The basilar membrane acts as a sort of Fourier analyser. It reacts to each frequency component of a sound, with the point of maximal





vibration near one end for low frequencies and near the other for high frequencies. Two simultaneous frequencies that are less than a critical bandwidth apart cause vibration patterns that interact on the membrane.

This is why it is difficult to hear individual tones in chord clusters with small pitch distances between adjacent notes.

The perceived pitch of a sound corresponds to its energy at integer multiples — harmonics — of a fundamental frequency. Two sounds containing harmonics within critical bandwidths make interference patterns on the basilar membrane, and produce a sense of dissonance. The frequency content of tones is also processed according to when neurons fire, and consonant and dissonant stimuli cause different types of firing patterns in auditory nerves.

Emotions arise in part through the ebb and flow of tension in music. Alternation between consonance and dissonance is a powerful device in this regard. Dissonance can be very beautiful, and resolution to consonance especially poignant. Composers can choose to ignore the fundamental physiological power of the consonance–dissonance relation and the information-processing power of discrete pitch and unequal interval sizes in scales. But by doing so, they create music that demands more of the listener because it lacks some of the most powerful physiological organizing principles of our nervous system.

Perception also depends on experience. During development, infants and children learn the pitch organization of their culture's music types and thereafter process music through the filter of this knowledge. Even musically untrained Western listeners acquire implicit knowledge of the Western major scale. They readily detect a wrong note that goes outside the scale on which a melody is based. They have

considerably more trouble detecting changes within the scale because these do not violate their implicit knowledge of which notes belong in the key. Infants under one, on the other hand, do not yet process music according to particular scales, and notice changes that violate major scale structure and changes that do not.

Experience counts

Harmonic structure (sequences of chords that follow each other according to syntactic rules) dominates Western music, but is relatively rare across musical systems. Without specific musical training, sensitivity to harmonic structure emerges in children only after about 5 years of age⁴.

Scale and harmonic structures depend on learning. By contrast with sensory consonance and dissonance, there is more flexibility in how they are perceived. They also show greater diversity across the world's musics. For example, different intervals are used in the Western major and minor scales, the pentatonic scale, and the many melodic modes (*rāgas*) used in Indian classical music. Whereas many traditions, such as *rāga* improvisations, employ a drone, or single pitch, over which the melody is played, fully developed harmonic syntax as in Western music is very rare. Our ability to learn different scales and harmonic structures gives composers considerable flexibility for experimentation that audiences can perceive and appreciate.

To recap: the spectral and temporal organization of music — its rhythm and pitch — derives from our biology. Neural constraints dictate that some musical structures are easier

to perceive and learn, giving rise to some near-universal features of music. Music is difficult to process when consonance and dissonance do not anchor the ebb and flow of tension, and when all pitches are equally prominent. Such music has no point from which to interpret pitch intervals. For many listeners, this level of difficulty is not enjoyable. Equally, music that is too simple and predictable can be boring.

The flexibility of our auditory system and its dependence on learning enables us to invent different musical structures, and allows musical tastes to change with familiarity and experience. There are notable examples of audience revolts at premier performances of works that seem tame to future generations — Stravinski's ballet *The Rite of Spring* caused a riot and Beethoven's third symphony was incomprehensible to reviewers. What has not changed recently is our evolutionary inheritance, the structure of our sensory organs, our basic encoding of information and our visceral responses to features of sound that unleash the emotional power of music in our lives. ■

Laurel Trainor is director of the McMaster Institute for Music and the Mind, and the Auditory Development Lab at McMaster University, 1280 Main Street West, Hamilton, Ontario L8S4L8, Canada.

1. Chang, E. F. & Merzenich, M. M. *Science* **300**, 498–502 (2003).
2. Phillips-Silver, J. & Trainor, L. J. *Science* **308**, 1430 (2005).
3. Trainor, L. J., Gao, X., Lei, J., Lehtovarara, K. & Harris, L. R. *Cortex* (in the press).
4. Trainor, L. J. *Dev. Psychobiol.* **46**, 262–278 (2005).

See <http://tinyurl.com/6e9mt8> for further reading.

D. PARKINS

NEWS & VIEWS

CLIMATE CHANGE

Hot questions of temperature bias

Chris E. Forest and Richard W. Reynolds

An unseen measurement bias has been identified in global records of sea surface temperature. The discrepancy will need correction, but will not affect conclusions about an overall warming trend.

On page 646 of this issue, Thompson *et al.*¹ take a fresh look at the global temperature record throughout the twentieth century, which is both a central test of climate models and *prima facie* evidence for man-made global warming. After filtering out large-scale natural effects, they uncover a large discontinuity in the data in 1945, and trace its source to a change in the instrumental bias in the sea surface temperature (SST) record that occurred around that time, and has not previously been adjusted. But how did this discrepancy come about? And how does it affect the credibility of the temperature record and, by extension, models of global warming in general?

In answering the first question, it is important to realize that SST and land-temperature observations are very different. The network of measuring stations on land is relatively fixed, not tending to wander in space. At a given station, thermometers are continually and reproducibly calibrated to give a continuous, benchmarked record. Changes at individual stations that might affect the measurements — if a station is moved, for example, or if the surrounding environment changes — can be corrected. Discovering additional unexpected biases in land temperatures would mean finding systematic changes in the whole observing system. Given the attention paid to these issues in, for example, the four assessment reports of the Intergovernmental Panel on Climate Change (IPCC) over the past two decades², as well as in the US Climate Change Science Program's final report on temperature trends in the lower atmosphere³, it is generally accepted that such a development is highly unlikely. The records represent a best possible estimate of the changes in land-surface temperatures.

For the oceans, the situation is different. Until the 1970s, SST observations were made entirely from ships. (After 1970, temperatures were also measured using moored and drifting buoys and, from the early 1980s, using satellites.) Different ships used different measurement methods over the years, each of which potentially had different biases. Some measurements were made by lowering uninsulated buckets over the ship's side; these tend to produce colder temperatures, owing to the effects

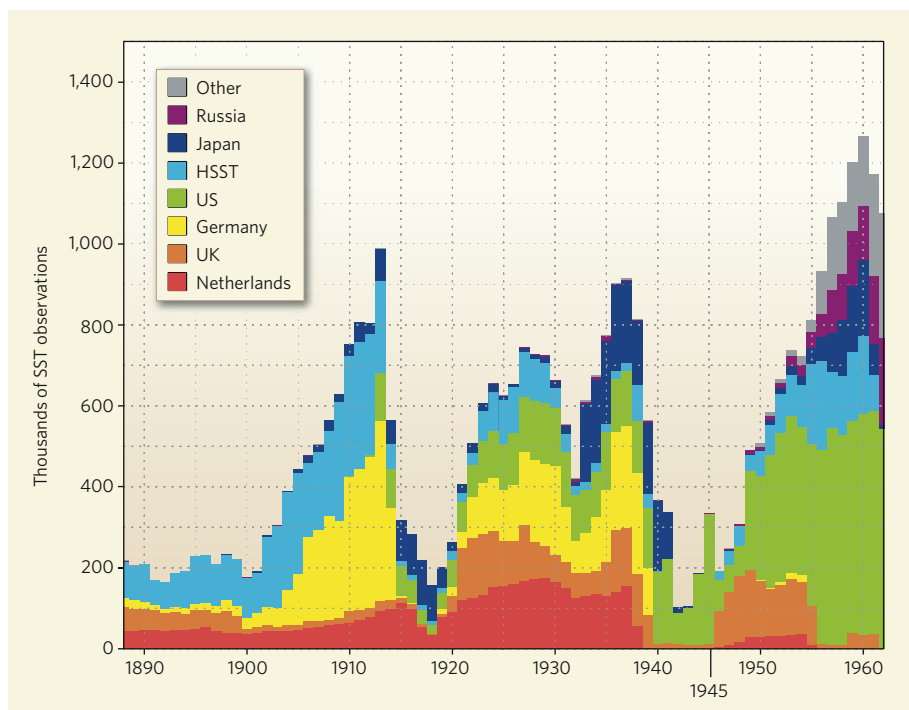


Figure 1 | Flying the flag. Sea surface temperature (SST) records from the International Comprehensive Ocean–Atmosphere Data Set^{15,16}, together with Historical Sea Surface Temperature (HSST) project data¹⁷, demonstrate a wide range of national origins — and, presumably, a similar range of slightly differing measurement protocols. Thompson *et al.*¹ identify a shift in the SST record in 1945 with the resumption of a significant programme of measurement by the United Kingdom at the end of the Second World War.

of evaporation once the bucket has left the water. Other measurements were taken at the inlet for the intake of water to cool the ship's engine; these are likely to be biased towards warmer temperatures because of heating from the engine-room.

Thus, although selected SST observations can be very accurate^{4,5}, corrections (generally of a few tenths of a degree Celsius) can be necessary depending on how the measurements were taken. Auxiliary information — metadata — about the measurement procedure is often missing, as reporting procedures differ from country to country, and the source of the data was not always recorded. Significant corrections to the global SST record had already been undertaken in response to recognized biases: an upwards movement of the data points before

1942 was a previous attempt to account for a shift from uninsulated-bucket measurements to engine-intake measurements.

Owing to the sparseness of the metadata, Thompson *et al.*¹ use a statistical method to identify where significant biases may occur. First, because short-term variability might mask these biases, they applied a simple filter to the data to remove two sources of significant natural climate variability: the El Niño–Southern Oscillation in the Pacific, and the interplay of land and ocean temperatures in the Northern Hemisphere. What remained were two filtered time series of the global mean temperatures for land and ocean. Comparing the two, several substantial jumps can be seen, but these are mainly the aftermath of large volcanic eruptions that ejected huge amounts

of material into the stratosphere, blocking out sunlight: Krakatoa (Indonesia) in 1883, Santa Maria (Guatemala) in 1902, Agung (Indonesia) in 1963–64, El Chichón (Mexico) in 1982, and Pinatubo (Philippines) in 1991. These jumps exist in both the land and ocean data. But one shift remained a puzzle: a significant drop in SSTs from 1945 to 1946 that was not replicated in the land data. This shift is also present, but not as obvious, in the unfiltered data. In an effort to explain this change, Thompson *et al.* looked to the meta-data — in particular, to the provenance of SST measurements from around 1945.

Figure 1 hints at the explanation; it shows the total number of SST observations from the various national temperature archives. What's striking is that both the relative fractions and the total numbers of observations vary considerably from year to year. These changes pervade the record, but unsurprisingly the two World Wars (1914–19 and 1939–45) represent the most significant shifts, both in source and in the total number of observations. And here, Thompson *et al.* suggest, we have the clue to the jump exposed in 1945: whereas during the preceding war years, 80% of measurements came from ships flying the US flag, a resumption of UK measurements at the end of the war saw their proportion jump to some 50%. At that time, unlike their American counterparts who took engine-intake measurements, the British relied primarily on uninsulated-bucket measurements.

So, what are the implications? Most immediately, a further correction to the global temperature series will be necessary, albeit of a magnitude yet to be assessed. There are many wider ramifications to consider, but one should be handled directly: should we doubt the rise in global mean temperatures during the twentieth century as a result of this or other hidden, and as yet undiscovered, biases in the record?

The answer is no. According to the filtering of natural variabilities that Thompson and colleagues have done, the only major discontinuity in SSTs is the one in 1945 (although other, insignificant shifts are dotted through the record). The shift from upwards-biased engine-intake measurements to downwards-biased bucket measurements demands a correction; naively speaking, temperatures between 1942 and 1945 would shift downwards by a magnitude of, say, 0.3 °C. Global warming would remain a reality — it would just be a bit more than previously thought.

How does this anticipated correction ripple through to climate models? Global mean surface temperatures are the most widely used data for evaluating the predictive capabilities of models on continental and larger scales^{6–8}. They are also crucial for evaluating two other principal uncertainties in climate predictions: the factors forcing climate change (primarily, levels of aerosol particles in the atmosphere) during the twentieth century and thus in the

future^{6–11}; and the rate of heat uptake from the atmosphere to the ocean (Fig. 1 in ref. 7).

The SST adjustment around 1945 is likely to have far-reaching implications for modelling in this period. One particularly striking example can be found in a figure prominently displayed on page 11 of the 18-page 'Summary for Policymakers'¹² of the IPCC's Fourth Assessment Report, produced last year. In this, the observed decadal mean temperatures of the 1940s — those that contain the anomalous 3–4-year interval dominated by (upwardly biased) US engine-room-intake measurements — are the only ones to lie above model predictions. Although we don't know exactly how the temperature record prior to 1946 will be affected by the awaited correction, it is a safe bet that temperatures in this particular decade will be lower.

The 1940s just happen to fall at the end of what seemed to be a warming trend from the 1910s that was followed by a weak decline in global temperatures into the 1960s. Climate modellers have explained the warming as a response to natural forcings¹³, and the cooling as due to an increase in tropospheric aerosols, principally sulphates, as a result of increased economic activity in the decades following the Second World War. This temporarily offset the effects of man-made warming. Data analysts, on the other hand, have considered the maximum in the 1940s to be the expression of a natural fluctuation¹⁴. In light of the new finding¹, each interpretation will need to be reconsidered — the first of many implications that will need to be explored. ■

Chris E. Forest is in the Joint Program on the Science and Policy of Global Change, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA. Richard W. Reynolds is in the US National Climatic Data Center, National Oceanic and Atmospheric Administration, 151 Patton Avenue, Asheville, North Carolina 28801, USA.
e-mails: ceforest@mit.edu;
richard.w.reynolds@noaa.gov

1. Thompson, D. W. J., Kennedy, J. J., Wallace, J. M. & Jones, P. D. *Nature* **453**, 646–649 (2008).
2. www.ipcc.ch/ipccreports/assessments-reports.htm
3. www.climate-science.gov/Library/sap/sap1-1/default.php
4. Kent, E. C. & Taylor, P. K. *J. Atmos. Ocean. Technol.* **23**, 464–475 (2006).
5. Kent, E. C. & Kaplan, A. J. *Atmos. Ocean. Technol.* **23**, 487–500 (2006).
6. Forest, C. E. *et al. Science* **295**, 113–117 (2002).
7. Stott, P. A. & Forest, C. E. *Phil. Trans. R. Soc. A* **365**, 2029–2052 (2007).
8. Knutti, R., Stocker, T. F., Joos, F. & Plattner, G.-K. *Clim. Dynam.* **21**, 257–272 (2003).
9. Andronova, N. G. & Schlesinger, M. E. *J. Geophys. Res.* **106**, 22605–22612 (2001).
10. Hegerl, G. C. *et al. in Climate Change 2007: The Physical Science Basis* (eds Solomon, S. *et al.*) 663–746 (Cambridge Univ. Press, 2007).
11. Meehl, G. A. *et al. in Climate Change 2007: The Physical Science Basis* (eds Solomon, S. *et al.*) 747–846 (Cambridge Univ. Press, 2007).
12. www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-spm.pdf (2007).
13. Stott, P. A. *et al. Science* **290**, 2133–2137 (2000).
14. Schlesinger, M. E. & Ramankutty, N. *Nature* **367**, 723–726 (1994).
15. Woodruff, S. D., Slutz, R. J., Jenne, R. L. & Steurer, P. M. *Bull. Am. Meteorol. Soc.* **68**, 1239–1250 (1987).
16. Kent, E. *et al. Bull. Am. Meteorol. Soc.* **88**, 559–564 (2007).
17. Slutz, R. J. *et al. Comprehensive Ocean–Atmosphere Data Set: Release 1* (NOAA Environmental Research Laboratories, Climate Research Program, Boulder, CO, 1985).

IMMUNOLOGY

Soothing intestinal sugars

Marika C. Kullberg

The gut is a new frontier in microbiology, offering many opportunities for innovative investigation. The finding of one such study is that intestinal inflammation in mice can be tamed by bacterial sugars.

The human 'gut flora' consists of between 300 and 1,000 microbial species, and some 10¹⁴ microorganisms in total (about ten times the number of cells of the human body). We usually live in harmony with these microbes, and would be less healthy without them. For example, they synthesize essential vitamins and amino acids, and also degrade otherwise indigestible plant material, as well as certain drugs and pollutants.

On page 620 of this issue, Mazmanian *et al.*¹ report that *Bacteroides fragilis*, a common bacterium of the lower gastrointestinal tract in mammals, can prevent intestinal inflammation in mice. Specifically, the authors show that polysaccharide A (PSA) of *B. fragilis* prevents gut inflammation induced by another

bacterium, *Helicobacter hepaticus*, or by the chemical compound TNBS (2,4,6-trinitrobenzene sulphonic acid).

This is an exciting finding, not least given that the incidence of human intestinal inflammation and inflammatory bowel disease has increased steadily in the Western world since the early 1950s. These conditions, which include Crohn's disease and ulcerative colitis, are believed to stem in part from inappropriate immune responses to the gut microbiota². In the healthy intestine, immune balance is regulated by different types of white blood cells called CD4⁺ T lymphocytes. These cells include CD4⁺ effector T lymphocytes (which help us fight pathogens by secreting various immune mediators called cytokines) and CD4⁺

regulatory T lymphocytes (which, through their production of the cytokines IL-10 and TGF- β , dampen the effector T cells when their action is no longer needed). When the balance between these types of T cell is disturbed, the immune response goes awry and intestinal inflammation occurs².

Evidence in support of the theory that bacteria trigger gut inflammation came from the discovery that mutations in NOD2, a host immune-cell receptor involved in detecting bacterial peptidoglycan, are associated with an increased risk of Crohn's disease^{3,4}. Furthermore, treatment with broad-spectrum antibiotics or probiotics (beneficial microbial species⁵) can improve health in patients with inflammatory bowel disease. Probiotics might inhibit the growth or invasion of pathogenic bacteria, or strengthen the gut-wall barrier. They may also stimulate the production of IgA (antibodies that are secreted into the gut lumen), and of IL-10 and TGF- β (ref. 5).

Mazmanian *et al.*¹ used two experimental approaches. In the first, colitis was induced in mutant (*Rag*^{-/-}) T-cell-deficient mice by infecting the animals with *H. hepaticus* and giving them CD4⁺ effector T cells. These effector cells start to respond to *H. hepaticus*, but in the absence of counterbalancing host regulatory T cells intestinal inflammation develops within a few weeks. The cytokine IL-23 plays a key role in this inflammatory response, and the disease is associated with a T_H1/T_H17 effector T-cell response to *H. hepaticus*⁶ (Fig. 1a). The second approach involves a chemical-induced colitis, in which the administration of TNBS to normal mice leads to acute inflammation within a few days.

In their new paper, Mazmanian *et al.*¹ demonstrate that giving mice *B. fragilis* at the same time as the colitis-inducing agents improves the animals' health. A mutant of *B. fragilis* that lacks PSA could not prevent inflammation, implicating PSA in helping to maintain immune balance. Finally, colitis did not develop when purified *B. fragilis* PSA was administered together with *H. hepaticus* and effector T cells to the *Rag*^{-/-} mice, or when this polysaccharide was given to the TNBS recipients. Together, these results provide direct evidence that PSA prevents intestinal inflammation in both model systems.

Bacteroides fragilis synthesizes at least eight distinct surface polysaccharides as part of its capsule, or coat⁷, PSA being the most abundant. Although CD4⁺ T cells normally recognize and respond to peptide fragments of proteins, PSA can be taken up by so-called antigen-presenting cells, such as dendritic cells, and presented to CD4⁺ T cells, resulting in T-cell activation⁷. Mazmanian *et al.*¹ show that IL-10 secreted by T cells is essential for PSA to protect against colitis. But it is not yet clear whether PSA has induced 'true' IL-10-secreting regulatory T cells in the mice protected from colitis. Thus, although the authors demonstrate that PSA cannot prevent colitis when T cells

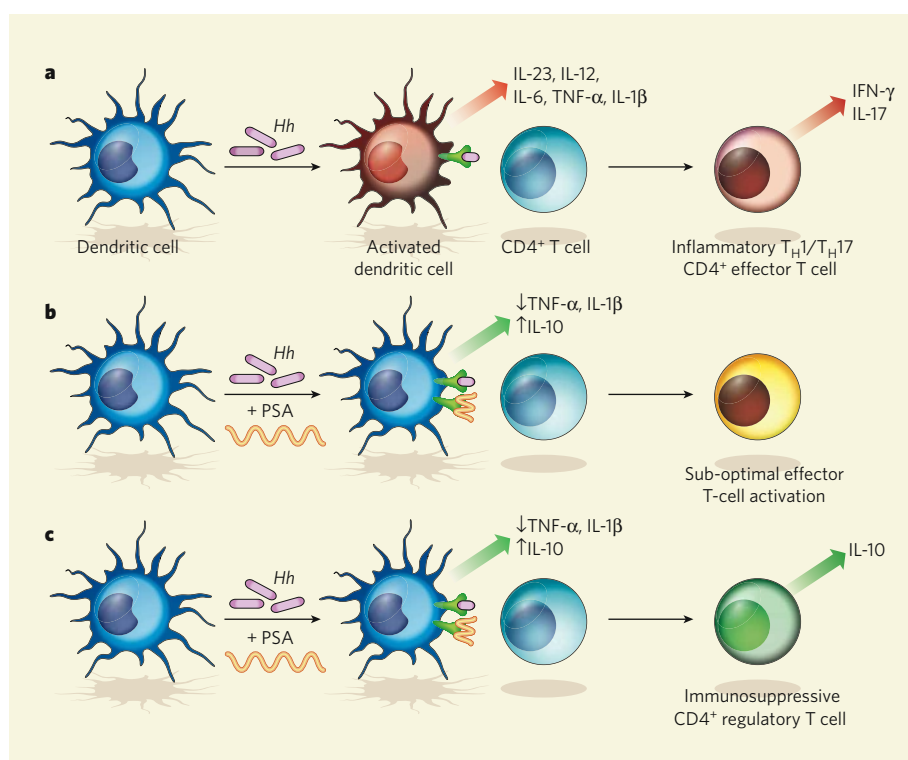


Figure 1 | Intestinal inflammation, and how *B. fragilis* PSA may prevent it in mice¹. **a**, In mice that develop colitis, dendritic cells activated by *H. hepaticus* (*Hh*) produce pro-inflammatory cytokines (such as IL-23, IL-12, IL-6, TNF- α and IL-1 β). They also present *Hh* antigen to CD4⁺ T cells that then differentiate into inflammatory T_H1/T_H17 CD4⁺ effector T cells; these cells are specific for *Hh* antigen, and produce the inflammatory cytokines IFN- γ and IL-17 (ref. 6). **b, c**, Dendritic cells jointly stimulated with *Hh* and *B. fragilis* PSA are less responsive, with reduced secretion of the pro-inflammatory cytokines TNF- α and IL-1 β , and enhanced production of anti-inflammatory IL-10 (ref. 1). They present both *Hh* antigen and PSA to CD4⁺ T cells, and may either (**b**) produce only sub-optimal effector T-cell activation or (**c**) induce the differentiation of IL-10-producing CD4⁺ regulatory T cells. The antigen specificity (*Hh* antigen and/or PSA) of such regulatory T cells is unknown.

cannot produce IL-10, this could be due to a more pathogenic nature of such IL-10-deficient effector T cells.

What about the molecular mechanism by which PSA prevents *H. hepaticus*-induced colitis? One possibility is that this sugar acts on dendritic cells, thereby altering their capacity to trigger an efficient effector-T-cell response (Fig. 1b). Support for this hypothesis comes from studies showing that T cells isolated from PSA-treated animals are hyporesponsive⁸.

Alternatively, is PSA inducing IL-10-secreting regulatory T cells (Fig. 1c)? Treatment with filamentous haemagglutinin (FHA) from *Bordetella pertussis*, the causative agent of whooping cough, can protect *Rag*^{-/-} mice from colitis induced by CD4⁺ effector T cells⁹. T cells isolated from these FHA-treated mice produced IL-10 *in vitro* following polyclonal stimulation⁹, supporting the model shown in Figure 1c. It will be interesting to see whether the same observation (that IL-10 is produced by T cells isolated from disease-protected mice) applies in the *H. hepaticus* colitis model in which PSA prevents intestinal inflammation.

Moreover, do CD4⁺ T cells isolated from disease-free mice infected with *B. fragilis* plus *H. hepaticus* respond to *B. fragilis* and/or *H. hepaticus* antigens? If so, what cytokines

do such bacterium-specific T cells produce? Has PSA skewed the T-cell immune response to *H. hepaticus* antigens away from a pro-inflammatory T_H1/T_H17-type response (Fig. 1a) towards IL-10 secretion (Fig. 1c)?

There is also the question of whether PSA can act against *H. hepaticus*-induced colitis not only when administered from the start of an experiment, but also when used to treat established disease. The long-term aim, of course, is to develop drugs to cure intestinal inflammation in humans. In this regard, treatment of patients suffering from inflammatory bowel disease with live eggs from the porcine whipworm, *Trichuris suis*, has produced promising results¹⁰. These effects are believed to be due to the induction by *T. suis* of regulatory T cells and factors such as IL-10, TGF- β and prostaglandin E2 that help maintain immune balance¹⁰. One view is that the apparent boons of modern life — antibiotics, vaccines and improved sanitation — have reduced the incidence of parasitic worms and other microbes, and therefore also of disease-protective molecules such as 'T. suis-like antigens' or PSA. That in turn may have altered the way our immune system responds to challenges, leading to the increased incidence of inflammatory diseases.

Ever-improving molecular techniques are providing tantalizing glimpses of the gut ecosystem¹¹. With the launch of the Human Microbiome Project¹², which plans to characterize the human microbiota and analyse its role in human health and disease, we are set to see considerable advances in understanding how host–microbial interactions may affect human health. When such information will translate into new therapeutic approaches is, however, anyone's guess.

Marika C. Kullberg is in the Immunology and Infection Unit, Department of Biology, University of York, and The Hull York Medical School, PO Box 373, York YO10 5YW, UK. e-mail: mk512@york.ac.uk

1. Mazmanian, S. K., Round, J. L. & Kasper, D. L. *Nature* **453**, 620–625 (2008).
2. Coombes, J. L., Robinson, N. J., Maloy, K. J., Uhlig, H. H. & Powrie, F. *Immunol. Rev.* **204**, 184–194 (2005).
3. Hugot, J.-P. *et al.* *Nature* **411**, 599–603 (2001).
4. Ogura, Y. *et al.* *Nature* **411**, 603–606 (2001).
5. Sartor, R. B. *Curr. Opin. Gastroenterol.* **21**, 44–50 (2004).
6. Kullberg, M. C. *et al.* *J. Exp. Med.* **203**, 2485–2494 (2006).
7. Mazmanian, S. K. & Kasper, D. L. *Nature Rev. Immunol.* **6**, 849–858 (2006).
8. Stingele, F. *et al.* *J. Immunol.* **172**, 1483–1490 (2004).
9. Braat, H. *et al.* *Gut* **56**, 351–357 (2007).
10. Elliott, D. E., Summers, R. W. & Weinstock, J. V. *Int. J. Parasitol.* **37**, 457–464 (2007).
11. Dethlefsen, L., McFall-Ngai, M. & Relman, D. A. *Nature* **449**, 811–818 (2007).
12. <http://nihroadmap.nih.gov/hmp>

See Editorial, page 563.

CANCER

Whispering sweet somethings

Thea Tlsty

That genetic mutations contribute to cancer is undisputed. What now emerges is that a cancer cell's microenvironment has a much stronger hand in the course a cancer takes than previously thought.

The goal of personalized medicine is to tailor a treatment to a specific cellular target at the appropriate stage of a disease, thus 'defusing' the disease process. Cancer is an example of the way in which multifaceted approaches to attaining this goal are emerging. We have come to appreciate that a tumour is a collection of diverse cells — cells carrying cancer-causing mutations and the cells of its immediate microenvironment — that act in concert towards disease progression^{1,2}. Three studies^{3–5} illustrate how these cells collude, and focus on the contribution of non-tumour cells.

Within tissues, epithelial cells are supported by a connective framework called the stroma, which itself consists of specific cells, such as fibroblasts, endothelial cells and immune cells, as well as the extracellular matrix. Physiological processes occurring in this compartment, for example the development of new blood vessels in response to oxygen shortage, and host immune responses, could dictate cancer progression.

Writing in *Nature Medicine*, Finak *et al.*³ set out to examine how gene-expression profiles in cells of the stroma are affected by cancer. Comparing morphologically normal and tumour stroma from the breast tissue of patients with breast cancer, they identify gene-expression patterns that are strongly associated with a specific outcome of the disease and that could be used as predictors of disease progression.

One specific predictor, a group of 26 genes that the authors call the stroma-derived prognostic predictor (SDPP), stratifies the risk of breast-cancer progression using molecular markers that are independent of — but add

power to — both standard clinical prognostic factors, such as the presence or absence of tumour cells in adjacent lymph nodes, and the more recently described⁶ predictors based on gene expression. SDPP identifies stromal subtypes that have gene-expression profiles relating to a good or poor outcome of breast cancer.

The clinical significance of work such as that of Finak *et al.* is twofold. First, discerning the subtleties of cell–cell interactions within the microenvironment of a malignant lesion (a localized, disease-associated change in a tissue) will indicate which particular therapy might be most effective for the specific biology of that tumour. Second, such insights could provide targets for developing new therapies. Finak *et al.* find that SDPP is not affected by treatment, suggesting that existing therapies do not target host responses that affect SDPP genes.

Reporting in *Proceedings of the National Academy of Sciences*, Postovit *et al.*⁴ use a contemporary approach to address the question of the stromal contribution to cancer malignancy. In this exciting study, the authors use an *in vitro* three-dimensional model that exposes cancer cells to the microenvironment to which human embryonic stem cells are normally exposed; they were hoping to identify conditions in the stroma that suppress malignant characteristics of cancer cells.

Stromal cells surrounding embryonic stem cells secrete a protein factor called Lefty, which inhibits the Nodal protein. Nodal, which during embryonic development prevents stem-cell differentiation, is abnormally expressed in human tumour cells, causing

malignancy⁷. Postovit and colleagues found that metastatic tumour cells do not express Lefty. Their results strongly support stromal regulation of malignancy and indicate that Lefty has a suppressive effect on cancer cells. The authors' work also suggests that factors secreted by the tumour stroma, and their derivatives, could be used as treatments to 'reprogramme' the differentiation of malignant cells, suppressing tumour development and growth.

Although modulating tumour properties in invasive cancers — as discussed in the Finak and Postovit papers^{3,4} — could reduce the associated morbidity and mortality, early diagnosis and prevention are even more effective means of preventing cancer-associated death. To address the clinical problems of cancer at these early stages, understanding the molecular processes underlying cancer initiation and progression is crucial. A paper by Hu *et al.*⁵ published in *Cancer Cell* addresses the mechanism of breast-cancer transition from a localized (*in situ*) lesion to an invasive form.

The authors used a cell line that, when injected into mice, mimics aspects of an early, non-malignant form of human breast cancer called ductal carcinoma *in situ* (DCIS)⁸, by forming non-invasive lesions in the animals' mammary gland. They next studied the role of myoepithelial cells in these lesions in suppressing the transition of DCIS to malignancy. (Myoepithelial cells separate the basement membrane of the duct from the epithelial cells that face the duct lumen.)

Hu and colleagues' functional analysis of cell-type-specific gene expression identified several pathways that could be essential for interactions between stromal fibroblast cells and myoepithelial cells in controlling the integrity of a tissue's basement membrane. These pathways, which modulate myoepithelial-cell differentiation, are mediated by essential signalling molecules such as TGF- β , Hedgehog, cell-adhesion molecules and the gene transcription factor p63. Malfunction of these signalling pathways leads to the loss of myoepithelial cells and subsequent invasion of the basement membrane by their adjoining epithelial cells, which respond to signals originating from fibroblasts. Determining whether the loss of myoepithelial cells is a cause or a consequence of the transition from *in situ* disease to invasive cancer will help to dictate therapeutic strategies.

Myoepithelial cells secrete a protein called maspin, which inhibits degradation of the extracellular matrix, an event thought to be essential for the transition from *in situ* cancer to an invasive form⁹. Moreover, this crucial tumour-suppressor protein is postulated to affect tissue invasion, programmed cell death and blood-vessel development. Hu and colleagues⁵ identified several extracellular-matrix metalloproteins that are implicated in cancer transition to the invasive state, but maspin is not one of them. Perhaps distinct subtypes of pre-malignant tumours use different pathways for the transition. Recent characterization of

DCIS tissue has indicated that it shows the same changes in gene-expression patterns as those seen in fully invasive tumours.

Understanding DCIS transition to invasive breast cancer is of tremendous importance. In the past two decades, the reported incidence of DCIS in the Western world has increased rapidly because of regular mammography screening. But only a small fraction of patients with DCIS will develop invasive disease or die of it¹⁰, as the vast majority have the tissue surgically removed and are then unlikely to develop subsequent tumours. This also means that additional treatment after surgery might be redundant. Nonetheless, after surgery, many women receive at least one of three regimens — radiation therapy, hormonal therapy or chemotherapy. At present, no criteria can consistently identify which women diagnosed with DCIS are most likely to benefit from these additional treatments. Identification of factors associated with subsequent invasive events could help classify women's individual risk for subsequent tumours and their response to therapies so as to avoid over- or under-treatment.

These three studies^{3–5} begin to deliver on the promise of basic research in characterizing the tumour microenvironment and the application of that knowledge in the clinic. Molecular markers in epithelial cells that predict which DCIS lesions will subsequently become invasive tumours are being found¹¹. By integrating such information with other data on the tissue microenvironment, researchers could identify additional molecular markers, thus improving prediction of future tumour formation and pointing the way to personalized treatment. The whisperings of molecular dialogue that go on between malignant cells and their microenvironment also hold information that can be used to categorize patients into those who need more, or less, aggressive therapy, identify patients for clinical trials, and develop new therapeutic approaches. These whisperings might even provide clinically important information before an invasive tumour can form.

Thea Tlsty is in the Department of Pathology, University of California, San Francisco, 513 Parnassus Avenue, San Francisco, California 94143-0511, USA.
e-mail: thea.tlsty@ucsf.edu

EVOLUTIONARY BIOLOGY

Sex ratios writ small

Jos. J. Schall

The evolutionary theory of sex ratios should apply to all creatures, both great and small. Experimental studies of the proportions of male to female sex cells of malaria parasites deliver cheering results.

Charles Darwin, the man of 'enlarged curiosity', was particularly curious about sex. He wondered, for example, why males and females are equally abundant in so many species in which males can mate with multiple females. Aren't males in surplus and a waste for such species? Darwin provided an answer, but was concerned primarily with human sex ratios. The question extends to even the single-celled protists, and on page 609 of this issue¹ Reece *et al.* revisit this venerable problem with a study of sex ratios in a protist that is both complex and lethal — the *Plasmodium* parasite that causes malaria.

Plasmodium prospers by replicating asexually within its vertebrate host, but also produces male and female gametocyte cells for transmission to a blood-feeding insect vector. Sex occurs within the vector, after each female gametocyte develops into a single female gamete, and each male yields several male gametes². Intuition suggests that a *Plasmodium* infection's transmission success into the vector would be greatest when just enough male gametocytes are present to mate with all the females. Female-biased sex ratios are indeed common, but an apparent surplus of male gametocytes is routinely seen in some *Plasmodium* species, and gametocyte sex ratio varies among and even within infections over time^{2–6}. Sex-ratio theory, a mainstay of modern evolutionary biology, offers explanations for these observations, but experimental verification has long been lacking.

Reece *et al.*¹ report that rodent malaria parasites follow sex-ratio theory quite well. Their elegant experiments show that each parasite clone shifts its ratio of male and female gametocytes according to the density of gametocytes in the blood, the fecundity of each male gametocyte and the likelihood of selfing (that is, union of male and female cells from the same clone). But *Plasmodium* also surprises with an additional talent — the parasite seems to detect kin and non-kin in the infection, and even the proportions of each.

Darwin provided a verbal explanation for the occurrence of equal proportions of males and females (Carl Düsing supplied the algebraic treatment a decade later)⁷. When sex ratio is biased, the less-common gender will have, on average, higher fitness, strictly because it will claim more offspring in the next generation. Mothers that produce offspring of the less-common gender would thus expect more 'grand offspring'. The equilibrium sex

ratio would be 1:1. Almost a century later, W. D. Hamilton recognized that this model holds only for outbred populations⁸. In a species that reproduces in patches in which sisters mate only with brothers, a mother's fitness depends on reducing competition among her sons for mates. Thus, just enough sons should be produced to mate with all the daughters. As the degree of mating between siblings declines within patches, the sex ratio should shift towards more equal representation of males and females. Humans show a 1:1 sex ratio because we are so well outbred.

Hamilton's model fits the life history of malaria parasites⁹. All mating of *Plasmodium* gametes occurs in a single blood meal within the vector. If an infection consists of a single genotype, or clone, of parasites, the optimal sex ratio for that clone would be one male gametocyte to f female gametocytes, where f is the fecundity of the male, or the number of gametes it produces. In mixed-clone infections, the optimal sex ratio for each genotype depends on the likelihood of selfing, and will shift appropriately towards more males.

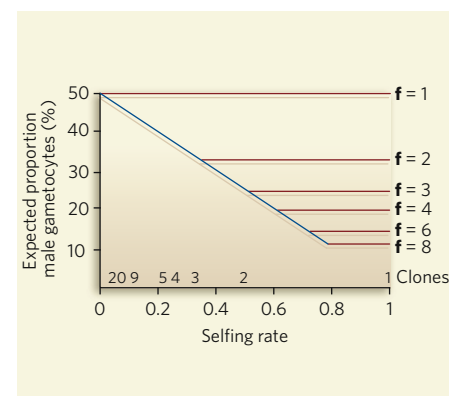


Figure 1 | A theory under test. Reece and colleagues¹ find that sex-ratio theory² predicts the proportion of malaria parasites' male gametocytes within a vertebrate host. The selfing rate depends on the number of genetically distinct clones and their proportions. Shown here are the expected rates of selfing with equal proportions of 1 to 20 clones. The fecundity of each male gametocyte (f) is the number of viable gametes produced per cell. With high fecundity ($f = 8$), and only one clone present, just enough males (11%) will be produced to mate with all the females within the insect vector. With low fecundity or many clones, the gametocyte sex ratio tilts towards higher production of males. (Figure redrawn from refs 2 and 4.)

- Weinberg, R. & Mihich, E. *Cancer Res.* **66**, 11550–11553 (2006).
- Tlsty, T. D. & Coussens, L. M. *Annu. Rev. Pathol.* **1**, 119–150 (2006).
- Finak, G. *et al.* *Nature Med.* **14**, 518–527 (2008).
- Postovit, L.-M. *et al.* *Proc. Natl Acad. Sci. USA* **105**, 4329–4334 (2008).
- Hu, M. *et al.* *Cancer Cell* **13**, 394–406 (2008).
- Paik, S. *et al.* *N. Engl. J. Med.* **351**, 2817–2826 (2004).
- Topczewska, J. M. *et al.* *Nature Med.* **12**, 925–932 (2006).
- Miller, F. R. J. *Mammary Gland Biol. Neoplasia* **5**, 379–391 (2000).
- Khalkhali-Ellis, Z. & Hendrix, M. J. *Cancer Res.* **67**, 3535–3559 (2007).
- Ketlikowski, K. *et al.* *J. Natl Cancer Inst.* **95**, 1692–1702 (2003).
- Gauthier, M. L. *et al.* *Cancer Cell* **12**, 479–491 (2007).

Hamilton's simple equations allow predictions of expected sex ratios (Fig. 1).

Two additions must be mentioned to account for wrinkles in the *Plasmodium* life cycle. First, the sex ratio of several *Plasmodium* species is female-biased early in the infection, but shifts towards more males as the infection ages^{5,6}. When the host mounts an immune attack against the parasite, carry-over of antibodies in the blood meal will kill many male gametes. Male fecundity will decline and more male gametocytes should be produced. Second, low-density infections may result in few gametocytes being transmitted, so male gametes cannot find a female. 'Fertility insurance' would then drive the production of more males¹⁰.

The Reece group¹ used well-characterized clones of *Plasmodium chabaudi* originally isolated from the natural host, African thicket rats (*Thamnomys*), and then inoculated into laboratory mice to initiate experimental infections. Real-time application of the polymerase chain reaction allowed quantification of specific genetic strains and precise measurement of sex ratio.

The authors found that single-clone infections were female-biased early on, but that over time the sex ratio shifted towards males. Single-clone infections should yield 11% male gametocytes early in an infection if $f = 8$ as per malariaiology lore. Four of the clones behaved according to theory. The two others produced more males, so we can predict the fecundity of these as 1 (the clone designated DK) and 4 (CR) (Fig. 1). Mixing all six clones should give 42% males, and this is just what was observed for the first six days of the infections. This outcome could be spurious if the DK clone dominates in infections (with its high male production), but this clone is known to be a poor competitor and to have low density in mixed infections. Mixing clones two-by-two, the expected result is 25% males, if both clones are equally abundant. But only one clone behaved as expected, with the others producing too few males.

However, Reece *et al.* determined the relative abundance of each clone, finding a negative correlation between the proportion of each parasite clone and its proportion of male gametocytes; when a clone predominated, it was more likely to self, and so produced fewer male cells. Finally, infections with a low density of gametocytes produced more males, even when only a single clone was present, which matches the expectations of fertility insurance.

These results should give cheer to fans of sex-ratio theory because the theory applies even for protist parasites dwelling within blood cells. Hamilton's equations are so simple, yet work so well. This is the real wonder of Reece and colleagues' study; it is as though this 'simple' parasite knows a little algebra.

Further questions have arisen, of course. There seems to be genetic variation for male fecundity (among isolates); why should this

be? How does the parasite recognize its own density in the host, and — even more vexing — how does it monitor the presence of kin versus non-kin in other blood cells? Finally, Reece and colleagues' experiments are a study in evolutionary ecology, but in this case the parasite and host have not coevolved, and the ecology is foreign. When a parasite of thicket rats enters a lab mouse, it meets a strange environment. Yet the protist follows the rules laid down in sex-ratio theory. Getting the gametocyte sex ratio right seems to be crucial for *Plasmodium*, no matter what host it visits. Once again, when dealing with sex, it seems that getting it right is all-important. ■

Jos. J. Schall is in the Department of Biology, University of Vermont, Burlington,

Vermont 05405, USA.

e-mail: jschall@zoo.uvm.edu

1. Reece, S. E., Drew, D. R. & Gardner, A. *Nature* **453**, 609–614 (2008).
2. Read, A. F. *et al.* in *Sex Ratios: Concepts and Research Methods* (ed. Hardy, I. C. W.) 314–332 (Cambridge Univ. Press, 2002).
3. Schall, J. J. *Adv. Parasitol.* **37**, 255–333 (1996).
4. Osgood, S. M. & Schall, J. J. *Parasitology* **128**, 23–29 (2004).
5. Paul, R. E. L., Coulson, T. N., Raibaud, A. & Brey, P. T. *Science* **287**, 128–131 (2000).
6. Paul, R. E. L., Brey, P. T. & Robert, V. *Trends Parasitol.* **18**, 32–38 (2002).
7. Edwards, A. W. F. *Am. Nat.* **151**, 564–569 (1998).
8. Hamilton, W. D. *Science* **156**, 477–488 (1967).
9. Read, A. F., Narara, A., Nee, S., Keymer, A. E. & Day, K. P. *Parasitology* **104**, 387–395 (1992).
10. West, S. A., Smith, T. G., Nee, S. & Read, A. F. *J. Parasitol.* **88**, 258–263 (2002).

MATHEMATICAL PHYSICS

Packings close and loose

Francesco Zamponi

What determines how grains such as sand pack together to fill a space? A thoroughgoing investigation of how geometry and friction interact in such systems is a step towards a more general understanding.

How should we arrange objects to pack them as tightly as possible, making best use of all the available space? Packing problems have long fascinated both physicists and mathematicians, but have proved surprisingly tough nuts to crack. Take the 'Kepler conjecture', for instance. It was in 1611 that Johannes Kepler first suggested that the densest packing of identical spheres is achieved by cubic (face-centred cubic) and hexagonal arrangements, with a packing fraction of 74%. Carl Friedrich Gauss produced the first partial proof of this in 1831. What might be a final proof was published only in 1998. It is a 'proof by exhaustion', reached using modern computing power to crunch wearisomely through an inordinate number of possible packing configurations — and its ultimate veracity is still being checked.

Sphere packings are extremely important, not only in condensed-matter physics¹, where they describe the favoured configurations adopted by crystals, but also in computer science and mathematics², where they pop up in problems related to group theory, number theory and error-correcting codes. On page 629 of this issue, Song, Wang and Makse³ take a significant stride towards a unified theory of a particular type of packing — not of the regular packings of the Kepler conjecture, but the random, amorphous packings that model the behaviour of everyday granular materials such as sand and nuts (Fig. 1).

When spherical grains are randomly thrown into a box and shaken, they form an amorphous arrangement with a packing fraction of 64%, significantly lower than the 74% of the densest

possible crystalline packing. Remarkably, this final density — the signature of 'random close packing' — was found to occur however the samples were prepared: whether by throwing grains into a box, shaking them and allowing them to settle; depositing them randomly around a disordered 'seed cluster'; slowly compressing a looser arrangement; and so on.

If small regions of regular, crystalline packing are created first, a random close packing can then be continuously compacted until a denser, entirely crystalline structure is obtained⁴. When looking at individual configurations, therefore, the density value 64% does not seem to have any special importance. Its relevance must instead be related to the statistical properties of an ensemble of packings produced by a given method. Is random close packing favoured for entropic reasons, such that there are just many more ways of jumbling grains up to form a random close packing than any other configuration? Is it a well-defined 'metastable' state that can persist for a considerable time? Or is it related to a hidden critical point, such that particularly large numbers of particles must be rearranged to change the density (a quality characterized by a large 'correlation length')? Many attempts have been made to achieve the statistical description of random close packings that such questions demand^{5–9}. These studies, supported by numerical simulations, revealed how important geometry, and in particular the network of particle contacts⁵, was in determining the density and other structural features of the final packing.

Random loose packings are related to

S. BADZ/PGA/GETTY IMAGES



PHOTOLIBRARY.COM

Figure 1 | Balls from regular to random. Johannes Kepler investigated the highest packing density of regular arrangements of spheres — here illustrated (left) by a face-centred-cubic stacking of golf balls. Less regular packings (right) are the subject of Song and colleagues' investigations³.

random close packings, but are even more elusive. They are obtained by letting spheres settle very gently¹⁰; the loosest stable packings that have been achieved have a packing fraction of about 55%, and friction is known to determine their stability. Is there a consistent statistical theory that can account for both close and loose random packings? Are the close and loose packings special points, or do similar stable configurations with packing fractions between 55% and 64% exist? Do these packings have critical properties such as a large correlation length^{5,11}? What is the relative importance of geometry and friction?

Song *et al.*³ provide answers to some of these questions. They develop a consistent, although approximate, mean-field theory of 'jammed' amorphous packings. A mean-field approach works by modelling the average interaction between bodies, thus making it the same for all the bodies in the packing. It is usually the first step towards any more sophisticated computation. The authors start by deriving a relation between the unoccupied space in any locality and the local geometrical coordination, which is defined as the average number of contacts per particle. The mean-field assumption means that some of the niceties of particle correlations are neglected, but the result agrees well with experimental data¹². The derived relation indicates that the packing fraction, which is directly related to the free volume, is determined solely by the geometrical coordination.

The authors go on to show that not all geometrical contacts carry a non-zero force. As a consequence, they introduce a mechanical coordination number, defined as the average number of contacts carrying a non-zero force. On the basis of numerical simulations, they assume that this number is a universal function of the friction coefficient, and is independent of the way the sample is prepared.

The emerging picture is thus of mechanical coordination determined uniquely by friction, and geometrical coordination related to density. Considerations of general stability require that both coordination numbers are somewhere between 4 and 6 (the mechanical coordination number is by definition smaller than the geometrical coordination number).

One then finds a collection of states satisfying these bounds, and draws a phase diagram by plotting two independent variables chosen from among the density, friction and the two coordination numbers of the states against each other. Song *et al.*³ choose density and mechanical coordination number. The final prediction is that if, for instance, friction is fixed (as it is in experiments), one can obtain packings with a whole range of densities (Fig. 2). The authors thus rationalize what is suggested by many experiments from the perspective of statistical mechanics.

An intriguing question immediately raised is how one might predict in which precise state a given preparation procedure will end. The authors follow a 20-year-old suggestion¹³ in introducing a variable they term compactivity. Compactivity is akin to an inverse pressure, in that it decreases as density increases, and Song *et al.* produce some numerical

evidence that, like thermodynamical pressure, it could be a 'state variable' that links different, seemingly independent, experimental control parameters.

Similar ideas have emerged in related contexts, such as the physics of glasses and problems of combinatorial optimization^{8,14}. In the latter example, researchers have tried to find a relationship between the behaviour of algorithms searching for solutions and the presence of transitions between different phases, and to identify state variables that could characterize these phases. Results so far have been contradictory, with some of them bringing into question the validity of this idea in situations beyond simple mean-field models¹⁵. But the perspective offered by work such as this, and that of Song *et al.*³ — with its promise of transforming complex problems of non-equilibrium dynamics into much simpler statistical problems — is too fascinating to abandon, and the wait for new results will not be long.

Francesco Zamponi is in the Laboratoire de Physique Théorique, Ecole Normale Supérieure, 24 Rue Lhomond, F-75231 Paris cedex 05, France. www.lpt.ens.fr/~zamponi

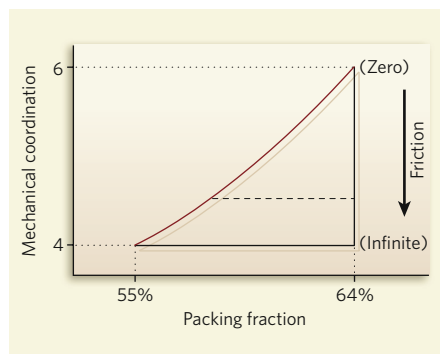


Figure 2 | Phases of packing. In Song and colleagues' phase diagram³, for a given friction (which determines the mechanical coordination) a range of possible densities exists (dashed line) — except at zero friction, where the density of random close packing, 64%, is a unique point. (The higher the mechanical coordination number, the lower the friction: friction blocks sliding modes and so fewer contacts are needed in order to make the packing stable.) Note that the geometrical coordination number for a given density is the maximum value of the mechanical coordination number. The 64% packing fraction of the random close packing corresponds to a geometric coordination number of 6; 55% (random loose packing) has a geometric coordination number of 4.

1. Torquato, S. *Random Heterogeneous Materials: Microstructure and Macroscopic Properties* (Springer, New York, 2006).
2. Conway, J. H. & Sloane, N. J. A. *Sphere Packings, Lattices and Groups* (Springer, New York, 1999).
3. Song, C., Wang, P. & Makse, H. A. *Nature* **453**, 629–632 (2008).
4. Torquato, S., Truskett, T. M. & Debenedetti, P. G. *Phys. Rev. Lett.* **84**, 2064–2067 (2000).
5. O'Hern, C. S., Silbert, L. E., Liu, A. J. & Nagel, S. R. *Phys. Rev. E* **68**, 011306 (2003).
6. Parisi, G. & Zamponi, F. *J. Chem. Phys.* **123**, 144501 (2005).
7. Parisi, G. & Zamponi, F. preprint at www.arxiv.org/abs/0802.2180 (2008).
8. Krzakala, F. & Kurchan, J. *Phys. Rev. E* **76**, 021122 (2007).
9. Kamien, R. D. & Liu, A. J. *Phys. Rev. Lett.* **99**, 155501 (2007).
10. Onoda, G. Y. & Liniger, E. G. *Phys. Rev. Lett.* **64**, 2727–2730 (1990).
11. Dauchot, O., Marty, G. & Biroli, G. *Phys. Rev. Lett.* **95**, 265701 (2005).
12. Aste, T., Saadatfar, M. & Senden, T. J. *J. Stat. Mech.* P07010 (2006).
13. Edwards, S. F. & Oakeshott, R. B. S. *Physica A* **157**, 1080–1090 (1989).
14. Altarelli, F., Monasson, R., Semerjian, G. & Zamponi, F. preprint at www.arxiv.org/abs/0802.1829 (2008).
15. Kurchan, J. in *Stealing the Gold: A Celebration of the Pioneering Physics of Sam Edwards* (eds Goldbart, P. M., Goldenfeld, N. & Sherrington, D.) 375–390 (Oxford Univ. Press, 2004).

Sex ratio adjustment and kin discrimination in malaria parasites

Sarah E. Reece^{1,2}, Damien R. Drew^{2,3} & Andy Gardner¹

Malaria parasites and related Apicomplexans are the causative agents of some of the most serious infectious diseases of humans, companion animals, livestock and wildlife. These parasites must undergo sexual reproduction to transmit from vertebrate hosts to vectors, and their sex ratios are consistently female-biased. Sex allocation theory, a cornerstone of evolutionary biology, is remarkably successful at explaining female-biased sex ratios in multicellular taxa, but has proved controversial when applied to malaria parasites. Here we show that, as predicted by theory, sex ratio is an important fitness-determining trait and *Plasmodium chabaudi* parasites adjust their sex allocation in response to the presence of unrelated conspecifics. This suggests that *P. chabaudi* parasites use kin discrimination to evaluate the genetic diversity of their infections, and they adjust their behaviour in response to environmental cues. Malaria parasites provide a novel way to test evolutionary theory, and support the generality and power of a darwinian approach.

The allocation of resources to male versus female offspring is one of the most well understood topics in evolutionary biology, and provides some of the best evidence for darwinian adaptation in the natural world^{1–5}. In many cases, simple theory can successfully predict when, why and by how much organisms should adjust their offspring sex ratio in response to changes in their environment. However, despite over a century of research effort directed at malaria (*Plasmodium*) and related Apicomplexan (including *Haemoproteus* and *Leucocytozoon*) blood parasites, evolutionary biologists have not been able to explain their sex ratios^{6–10}. These parasites replicate asexually, and a proportion of these asexually produced parasites develop into sexual stages, termed gametocytes. When taken up in an insect vector's blood meal, gametocytes differentiate into gametes and mate. In *Plasmodium* species, male gametocytes can differentiate into a maximum of eight gametes and each female differentiates into a single gamete. Sex allocation in *Plasmodium* is consistently female-biased when measured across populations, but sex ratios vary extensively when measured throughout individual infections^{11–15}. Sex is crucial for transmission to vectors, and so there is a drive to identify drugs and develop vaccines that block this process. Yet remarkably little is understood as to why such female-biased and variable sex allocation strategies have evolved.

Hamilton's theory of 'local mate competition' (LMC) predicts that female-biased sex allocation is favoured when genetically related males compete for mates; because males can each fertilize more than one female, an equal sex ratio would result in a wasteful surfeit of male gametes⁴. For malaria parasites, LMC predicts that the unbeatable sex allocation strategy (investment into males relative to females) depends on the inbreeding rate, and is given by the equation $z^* = (1 - f)/2$, where z^* is the proportional allocation into males versus females and f is Wright's coefficient of inbreeding^{16–19}. When inbreeding does not occur, owing to a large number of genotypes represented in mating groups (no inbreeding; $f = 0$), the unbeatable strategy is to produce an equal number of males and females²⁰ ($z^* = 1/2$). Conversely, when only one or a small number of genotypes are present (inbreeding; $f > 0$), female-biased sex allocation is favoured⁴ ($z^* < 1/2$).

LMC is one of the most successful theories in evolutionary biology and its explanatory power has been demonstrated in a variety of taxa²¹ including plants, snakes, insects, mites, worms and fish, but the application of LMC theory to Apicomplexan parasites has proved controversial and inconclusive. There is considerable variation in the inbreeding rate (as allowed by the number of multiple infections) experienced by parasites within and across species of Apicomplexa^{22–24}. LMC is supported by correlations between observed sex ratio and inbreeding rate across several Apicomplexan taxa, with female-biased sex ratios occurring in populations with higher rates of inbreeding^{19,25,26}. However, contradictory data are as numerous: (1) sex ratios in populations of *Haemoproteus* bird parasites do not correlate with the inferred genetic diversity of their infections and are consistently less female biased than expected⁹; (2) LMC theory cannot explain the considerable variation in sex ratios observed during experimental infections^{8,19,26–28}; (3) the inbreeding rate is not the only factor shaping sex allocation, because *Plasmodium* parasites facultatively alter their sex ratio in response to changes in host anaemia^{11,13}; and (4) successful transmission to vectors appears unrelated to mating group sex ratio^{29,30}.

Failure to understand sex allocation in malaria parasites poses problems for both medical science and evolutionary biology. If an evolutionary framework cannot explain a relatively simple trait like parasite sex allocation, there is no reason to believe that it can be usefully applied to more complex traits such as virulence. Also, organisms such as parasites and microbes present a novel and independent test for the explanatory power and generality of an evolutionary theory that has been largely developed to explain the biology of metazoan taxa such as insects, birds and mammals³¹. Explicitly testing whether sex allocation in malaria parasites is shaped by LMC has not previously been possible, because the required experimental methods and techniques have only just become available.

First, we test the basic assumptions of LMC as applied to malaria parasites. We test whether females are the more limiting sex for mating group productivity and whether the relationship between sex ratio and fitness varies in the manner predicted by theory^{1,4}. Even though LMC has been so successfully applied to other taxa, this

¹Institute of Evolutionary Biology, ²Institute of Immunology and Infection Research, Ashworth Laboratories, School of Biological Science, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK. ³The Walter and Eliza Hall Institute of Medical Research Biotechnology Centre, Research Avenue, Bundoora, Victoria 3086, Australia.

fundamental assumption has yet to be properly tested in any species. Such data are lacking owing to the difficulties in manipulating this trait independently of confounding variables, and the very success of LMC theory has suggested that such a test is unnecessary. We also test whether there is genetic variation for patterns of within-infection sex allocation in malaria parasites, and examine whether these patterns follow the predictions of theory^{32,33}. Second, we test the key predictions of LMC by manipulating the number of genotypes present in infections (inbreeding rate) and investigate whether focal genotypes adjust their sex allocation strategy in response^{13,26,34,35}. Facultative sex-ratio adjustment in response to genetic diversity would confirm that sex allocation strategies in malaria parasites are as sophisticated as those observed in multicellular taxa, that they can be explained by evolutionary theory, and that malaria parasites can discriminate kin from non-kin.

Fitness consequences of sex allocation

We used two genetically transformed lines of the rodent malaria *Plasmodium berghei* to examine the fitness consequences of sex ratio variation^{36,37}. One parasite line (*pb48/45-ko*) cannot express genes essential for male-gamete fertility, so it is limited to female function; the other (*pb47-ko*) cannot express genes essential for female-gamete fertility, so it has only male function. Otherwise, both lines are genetically identical, exhibit normal infection dynamics and their gametocyte and gamete production is the same as their wild-type ancestor. By mixing parasites from these lines in different proportions and using *in vitro* fertilization culture methods, we directly manipulated mating-group sex ratio (proportion of male gametocytes) and measured the resulting reproductive success as the number of zygotes (ookinetes) produced³⁸. As predicted by theory, mating success is maximized at intermediate sex ratios, indicating that sex allocation in malaria parasites is likely to be under stabilizing selection (Fig. 1: linear term, $F_{1,90} = 40.97$, $P < 0.0001$; and quadratic term, $F_{1,90} = 13.81$, $P = 0.004$). These data also support LMC theory, as mating group reproductive success is maximized at female-biased sex ratios (maximum 33% males: 95% confidence index (CI), 20–39%; see Supplementary Information). We estimate that each male produces an average of 2.03 viable gametes (95% CI, 1.56–4.00), revealing that, on average, male gametocytes fail to realize their potential fecundity of eight gametes and will become limiting to mating group productivity at extremely female-biased sex ratios. Owing to biological constraints, difficulty controlling for other relevant trade-offs, and genotype-by-environment interactions, mapping fitness across the range of sex allocation strategies has not

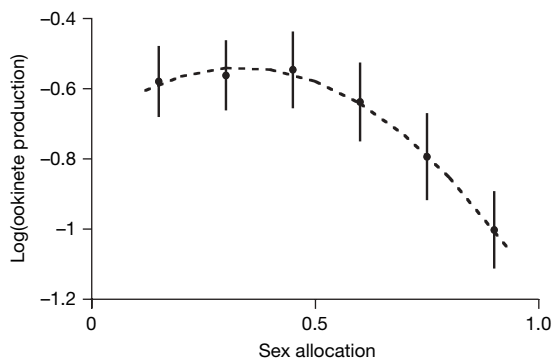


Figure 1 | The fitness consequences of sex ratio variation. The relationship between sex allocation (given as proportion male) and fitness (given as log(ookinete production)) varies in the manner predicted by theory, revealing that this life-history trait of malaria parasites is important and under selection. Shown is log-transformed mean ookinetes ($\times 10^6 \text{ ml}^{-1}$) produced in 19 cross-factored sets of *P. berghei* cultures spanning 0–100% males ($R^2 = 0.77$); dashed line is the fitted relationship. As expected, no ookinetes were produced in our control groups of 0% and 100% males, and we excluded these data from our analysis. Error bars, \pm s.e.m.

previously been possible for malaria parasites or any other taxa. Therefore, these data are the first to unequivocally support a fundamental assumption of sex allocation as applied to any species.

Genetic variation for sex allocation

We then used wild-type clonal genotypes of the rodent malaria *P. chabaudi* to test, for the first time, whether there is within-species genetic variation for patterns of sex allocation in malaria parasites. We sequenced six *P. chabaudi* genotypes and determined that our recently developed quantitative reverse-transcription PCR assays³⁹ for measuring sex ratios could be applied to all of them (see Supplementary Information). We initiated replicate controlled infections of each of these genotypes and measured infection parameters daily throughout the acute phase. We found that our six genotypes followed four significantly different sex allocation patterns throughout their infections (Fig. 2: $\chi^2_{55} = 159.55$, $P < 0.0001$; see Supplementary Table 1). Such genetic variation is required for selection to act, but could itself be adaptive and reflect differences in the number of gametes produced by males of different genotypes, a parameter which may also vary throughout infections^{30,32}. If this is the case, then we can estimate the fecundity of male gametocytes from sex ratio (see Supplementary Information). In addition to controlling for variation across days, our analyses controlled for infection parameters (including virulence, anaemia, asexual parasite and gametocyte density) that could influence sex ratio and confound differences between the genotypes. Only parasite density remained in the minimal model, which correlated positively with sex ratio ($\chi^2_1 = 5.25$; $P = 0.022$; slope = $(27.35 \pm 13.41) \times 10^6 \text{ ml}^{-1}$).

Sex ratio variation during infections

In addition to the presence of co-infecting genotypes, malaria parasites are predicted to allocate sex in response to factors that compromise their fertilization success in blood meals^{11,26,32,33}. Specifically, a lower female-bias is predicted for a given inbreeding rate when the fertility of males is reduced by factors such as low gametocyte density, host anaemia and immunity. There are two broad reasons why these factors are expected to reduce fertilization success. First, when gametocyte density is low and/or hosts are anaemic, there is a stochastic risk of too few male gametocytes being taken up in blood meals to fertilize the females. Second, if the appearance of host factors reduces the ability of male gametocytes to produce viable gametes, there will

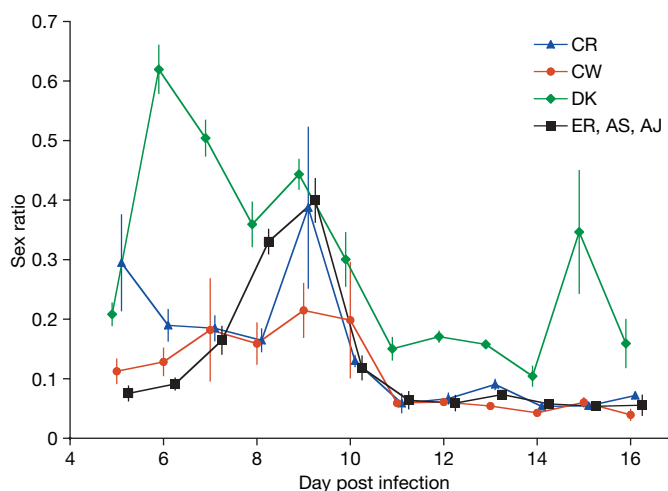


Figure 2 | Genetic variation in patterns of sex allocation. *P. chabaudi* genotypes exhibit significant genetic variation in the sex ratios produced throughout their infections. Genotypes DK, CW and CR all followed significantly different sex allocation patterns but AS, AJ and ER could be grouped together. Here, and in Figs 4–6, sex ratio is given as proportion male. The means from 30 independent infections are presented and the x-axis is jittered for clarity. Error bars, \pm s.e.m.

not be enough male gametes in the blood meal to fertilize the females. When few gametocytes are able to interact in a blood meal and/or male gamete production is low, parasites are expected to increase their investment in male gametocytes to ensure their females are fertilized. Data from our single genotype infections support these predictions, and show that throughout infections sex ratios correlate negatively with red blood cell density ($\chi^2_3 = 15.86$; $P = 0.0012$) and gametocyte density ($\chi^2_3 = 22.11$; $P < 0.0001$), but positively with parasite density (which is related to the strength of host immune responses; $\chi^2_3 = 35.35$; $P < 0.0001$), and that there is genetic variation for these patterns (Fig. 3 and Supplementary Table 2).

Genetic diversity and sex allocation

Having verified that the key assumptions of sex allocation theory are appropriate to the biology of malaria parasites, we tested whether they can facultatively respond to LMC. We compared sex ratios produced by our six single-genotype infections to sex ratios in mixed-infections consisting of all six genotypes. Our above analyses support the prediction that parasite sex ratios are influenced by other parameters^{11,26,32,33}, so we predicted that any response to the presence of co-infecting genotypes is most likely to be detected before these

factors exert their confounding influences. For all of our treatment groups, day 8 post infection (PI) was the modal day of peak parasite density, and parasite density was subsequently reduced by immune factors, anaemia and competition between genotypes. Therefore, we restricted our analyses to the period between infections becoming patent and reaching their peaks (days 5–7 PI). As predicted by LMC, sex ratios in six-genotype infections were less female-biased than in single-genotype infections, and the magnitude of this effect varied through infections (Fig. 4: $\chi^2_2 = 19.93$, $P < 0.0001$; see Supplementary Table 3). As with the inbreeding rate and sex ratio, we found a significant negative correlation between gametocyte density and sex ratio ($\chi^2_1 = 11.94$; $P < 0.0001$; slope = $(-28.20 \pm 8.18) \times 10^{-3} \text{ ml}^{-1}$). Using the model of ref. 25, and assuming that male gametocytes produce up to 8 gametes on average, the predicted evolutionary stable sex ratios for single- and six-genotype infections are 0.11 (proportion male) and 0.42 respectively. At the start of infections, our data support these predictions: on days 5 and 6 PI, sex ratios were respectively 0.14 ± 0.02 and 0.20 ± 0.04 in the single-genotype infections, and 0.46 ± 0.09 and 0.42 ± 0.07 in the six-genotype infections.

Sex allocation in focal genotypes

The best way to test whether parasites respond to LMC is to test whether focal genotypes alter their sex allocation in response to the presence of co-infecting genotypes. We were able to examine this for the three *P. chabaudi* genotypes AJ, AS and ER in single- and double-genotype infections. Our quantitative PCR assays could distinguish genotype AJ from genotypes AS and ER but not between AS and ER (see Supplementary Information). Therefore, we tested whether AS and ER independently altered their sex allocation when co-infecting with AJ, and whether AJ altered its sex allocation when co-infecting with either AS or ER. As we could confirm that parasites from both genotypes were present throughout the mixed infections, we were able to analyse sex ratios when parasites are in the growth phase (parasite density increasing) and post-peak phase (parasite density declining) of their infections. We split our analyses into the pre-peak period (days 5–7 PI) and the post-peak phase (days 8–12 PI). Because sex ratios produced by AJ when co-infecting with AS were not significantly different to those produced when co-infecting with ER (growth phase; $\chi^2_1 = 1.66$, $P = 0.198$; post-peak phase; $\chi^2_1 = 0.25$, $P = 0.617$), we grouped these AJ data. In the growth phase of their infections, both AJ and ER increased their investment in males when co-infecting with another genotype, as predicted by LMC, but AS did not (Fig. 5; $\chi^2_2 = 9.98$, $P = 0.007$; see Supplementary Table 4a). Sex ratio adjustment was greater in AJ than ER and all other infection parameters fitted in the maximal model were non-significant. In contrast, during the post-peak phase only AJ increased investment in males when infecting with a second genotype; AS decreased

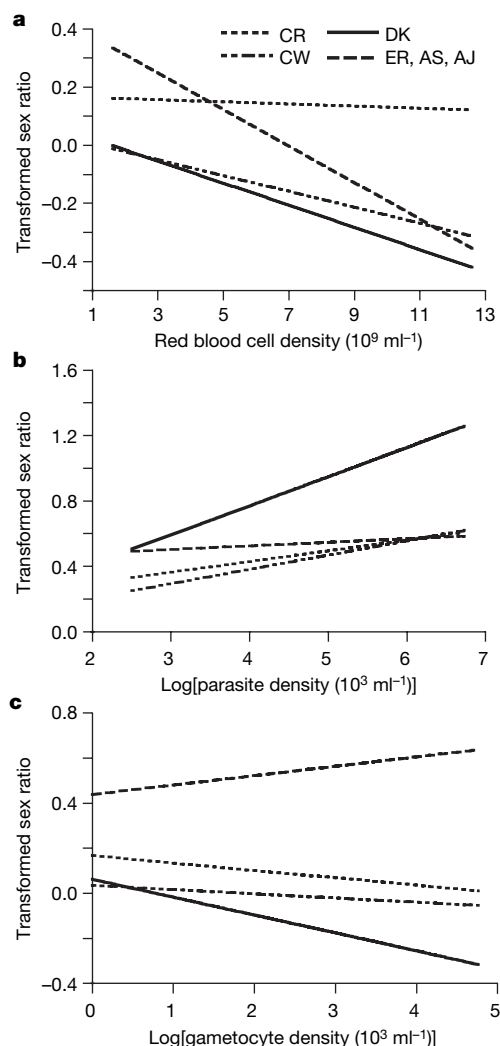


Figure 3 | Explaining sex ratio variation throughout infections. Sex ratios (arcsin square-root transformed) of *P. chabaudi* correlate with the density of: **a**, red blood cells; **b**, parasites; and **c**, gametocytes. Lines are fitted from the estimates predicted by the minimal model using infection parameters observed 48 h before sex ratios (see Supplementary Information). Genotypes are grouped according to the four different sex ratio patterns followed throughout 30 independent infections (Fig. 2).

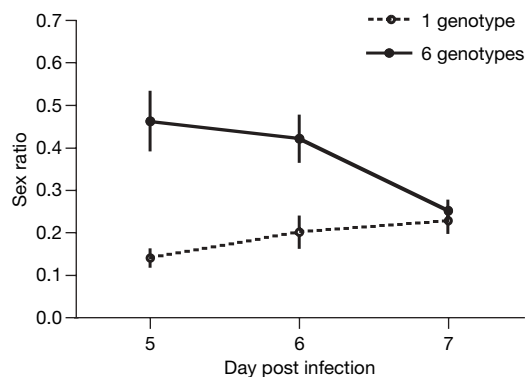


Figure 4 | Sex ratio varies with the genetic diversity of *P. chabaudi* infections. Infections with six genotypes produced significantly less female biased sex ratios than those with one genotype, but only at the start of infections. Means are presented from 40 independent infections. Error bars, \pm s.e.m.

investment in males and sex ratios were not significantly altered in ER (Fig. 6; $\chi^2_2 = 27.33$, $P < 0.0001$; see Supplementary Table 5a). All other infection parameters fitted in the maximal model were non-significant apart from red blood cell density, which correlated positively with sex ratio ($\chi^2_1 = 6.89$, $P = 0.009$; slope = $(32.91 \pm 13.06) \times 10^6 \text{ ml}^{-1}$).

Facultative sex allocation in response to infection genetic diversity of genotypes ER and AJ (when in two different co-infection scenarios) provides unequivocal support for LMC. Given this result, we extended our analysis to test whether sex allocation of focal genotypes correlates with their relative representation in infections. Theory predicts that genotypes making a small relative contribution of gametocytes to the mating group should invest in more males than when making a relatively large contribution⁴⁰. Using our co-infection data, we show that there is a negative correlation between sex ratio and the proportion of gametocytes contributed by focal genotypes during the pre-peak period ($\chi^2_1 = 16.08$, $P < 0.0001$; slope = -0.52 ± 0.12 ; see Supplementary Table 4b), but not during the post-peak period ($\chi^2_1 = 2.68$, $P = 0.102$; Supplementary Table 5b). Using proportional representation within the mating group as a proxy of inbreeding rate, LMC theory predicts a sex ratio of $z^* = (1 - f)/2$, and hence a slope of $dz^*/df = -0.5$; thus, the pre-peak data are quantitatively consistent with LMC theory. These data support the possibility that genotypes

may be able to infer their own relative frequency in infections. It is not clear why genotype AS did not produce less female-biased sex ratios in response to the presence of a co-infecting genotype. However, this genotype is substantially less virulent than genotypes AJ and ER. The least virulent genotype in our panel is DK⁴¹, and this genotype produced the least female-biased sex ratios in single infections. Virulence is a life-history trait unique to infectious organisms, and could influence reproductive strategies in malaria parasites. Understanding how malaria parasites evaluate the genetic diversity of their infections may also explain why genotype AS did not behave as predicted by LMC.

Discussion

We have tested and confirmed the assumptions and predictions of sex allocation theory in malaria parasites. The female bias typical of single-genotype infections declines towards an equal sex ratio in multiple infections. We also show that: (1) sex ratio is important and transmission studies must consider this trait alongside gametocyte density; (2) mating group productivity decreases at extremely female-biased sex ratios because male gametes become limiting; (3) within-infection sex ratio patterns are explained by variation in anaemia and parasite densities (patterns for which there is significant genetic variation); and (4) female-biased sex allocation in response to LMC decreases as infections progress and parasite densities decline due to competition, anaemia and the host's immune response. These data support 'fertility insurance theory', which predicts that malaria parasites should adjust their reproductive strategies to maximize transmission opportunities throughout their infections^{11,26,32,33}. Furthermore, our data do not support the hypothesis that malaria

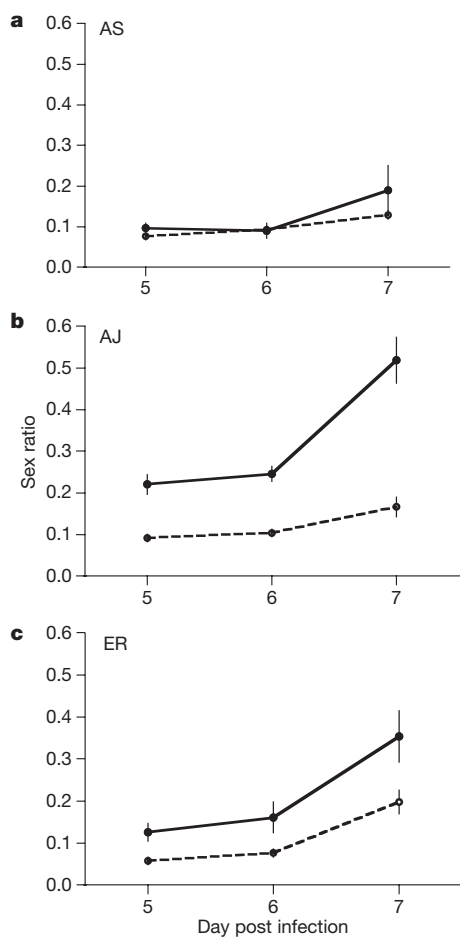


Figure 5 | Sex ratios of focal genotypes during the growth phase of infections. Mean sex ratios throughout the growth phase of infections for *P. chabaudi* focal genotypes when alone (dashed line) and co-infecting with a second genotype (solid line). **a**, AS; **b**, AJ; **c**, ER. Sex ratios of genotypes AS and ER could be distinguished from sex ratios of AJ but AS and ER could not be distinguished from each other (see Supplementary Information). Sex ratios produced by AJ when co-infecting with AS or ER were not significantly different so these infections are grouped. We followed 5 independent infections for each genotype combination. Error bars, \pm s.e.m.

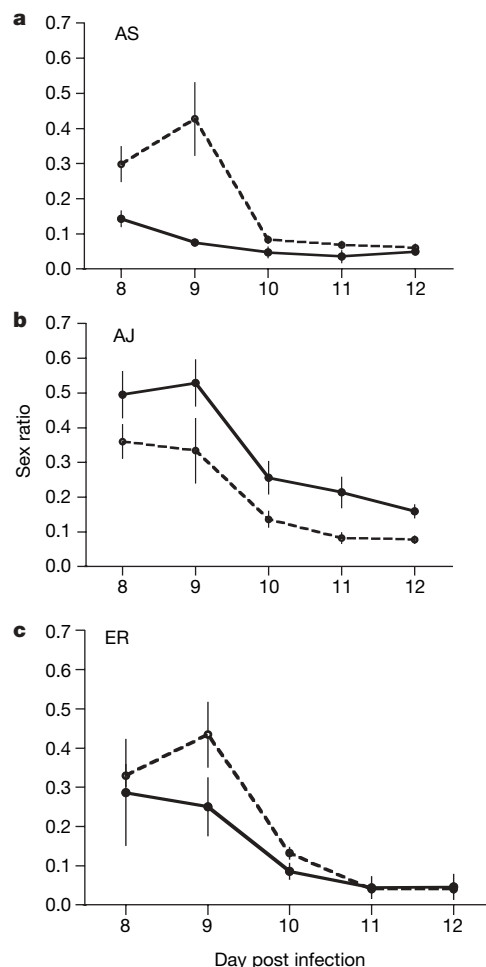


Figure 6 | Sex ratios of focal genotypes during the post-peak phase of infections. As Fig. 5 but for the post-peak phase.

parasites preferentially mate with clone-mates, regardless of the genetic make-up of their mating groups⁴². We have shown that an evolutionary approach can be successfully applied to sex allocation in malaria parasites, suggesting that researchers in pure and applied fields can be optimistic about understanding more complex parasite life-history traits, such as virulence, that are of such economic, medical and veterinary importance⁴³.

As infections consist of billions of parasites, the ability to distinguish clone-mates from unrelated conspecifics requires kin discrimination^{44,45}. This discovery raises important questions. Does within-infection relatedness influence competitive interactions between genotypes and shape patterns of virulence^{46–48}? Low relatedness between co-infecting malaria parasites is expected to favour the most virulent competitors, but related parasites could also cooperate to facilitate growth or immune evasion and result in more virulent infections. Apoptosis has recently been described in protozoan parasites⁴⁹, and kin discrimination supports the possibility that this is actually an altruistic behaviour. Whether malaria parasites are using indirect (environmental) or direct (genetic) cues to discriminate kin must now be investigated. Given that *P. chabaudi* parasites are able to detect the presence of conspecifics in a non-natural host environment (see Supplementary Information), they may be using direct cues rather than changes to their environment. Genetic kin discrimination is rare, but can evolve in situations where marker diversity will be maintained by extrinsic processes^{45,50}. Host–parasite interactions may provide a strong enough source of balancing selection; does pressure to evade immune recognition also enable malaria parasites to employ their sophisticated social behaviour?

METHODS SUMMARY

We used recently developed genetic modification and molecular techniques to carry out these experiments. The fitness experiment used *P. berghei* lines from the same genetic background in which genes essential to either male or female function have been knocked out. Methods for *in vitro* culture of mating groups are based on previously published protocols³⁸. We measured reproductive success by counting ookinetes (zygotes at 18 h of age). Morphological identification is straightforward, as unfertilized female gametes and gametocytes are spherical and ookinetes are elongated banana-shaped cells. We used *P. chabaudi* for our sex ratio experiments for two reasons. First, natural infections vary in genetic diversity and second, a bank of distinct and characterized clonal genotypes are available for this species. We used the genotypes AS, AJ, ER, CR, CW and DK (WHO Registry of Standard Malaria Parasites, The University of Edinburgh) and our experimental design was based on previously published competition treatments. We used previously developed quantitative reverse-transcription PCR assays for a gametocyte-specific and a male-specific gene to calculate sex ratios (proportion of gametocytes that are male). We used 6–8 week old male MF1 mice (in-house supplier, University of Edinburgh). All mice received an intraperitoneal (i.p.) inoculation of parasitized red blood cells in a 0.1 µl dose as previously described¹¹. Mice were housed at 21 °C with a 12 h light cycle, and maintained on a diet of SDS41B food pellets (Harlan Scientific) and 0.05% PABA supplemented drinking water to enhance parasite growth.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 29 February; accepted 28 March 2008.

- Charnov, E. L. *The Theory of Sex Allocation* (Princeton Univ. Press, Princeton, 1982).
- Frank, S. A. Sex allocation theory for birds and mammals. *Annu. Rev. Ecol. Syst.* **21**, 13–55 (1990).
- Frank, S. A. A touchstone in the study of adaptation. *Evolution Int. J. Org. Evolution* **56**, 2561–2564 (2002).
- Hamilton, W. D. Extraordinary sex ratios. *Science* **156**, 477–488 (1967).
- Hardy, I. C. W. *Sex Ratios: Concepts and Research Methods* (Cambridge Univ. Press, Cambridge, UK, 2002).
- Ferguson, D. J. P. *Toxoplasma gondii* and sex: Essential or optional extra. *Trends Parasitol.* **18**, 355–359 (2002).
- Ferguson, D. J. P. More on *Toxoplasma gondii*, sex and premature rejection. *Trends Parasitol.* **19**, 157–158 (2003).
- Paul, R. E. L., Arie, F. & Robert, V. The evolutionary ecology of *Plasmodium*. *Ecol. Lett.* **6**, 866–880 (2003).
- Shutler, D., Bennett, G. F. & Mullie, A. Sex proportions of *Haemoproteus* blood parasites and local mate competition. *Proc. Natl Acad. Sci. USA* **92**, 6748–6752 (1995).
- West, S. A., Reece, S. E. & Read, A. F. *Toxoplasma gondii*, sex and premature rejection. *Trends Parasitol.* **19**, 155–157 (2003).
- Reece, S. E., Duncan, A. B., West, S. A. & Read, A. F. Host cell preference and variable transmission strategies in malaria parasites. *Proc. R. Soc. Lond. B* **272**, 511–517 (2005).
- Robert, V. *et al.* Sex ratio of *Plasmodium falciparum* gametocytes in inhabitants of Dielmo, Senegal. *Parasitology* **127**, 1–8 (2003).
- Paul, R. E. L., Coulson, T. N., Raibaud, A. & Brey, P. T. Sex determination in malaria parasites. *Science* **287**, 128–131 (2000).
- Paul, R. E. L., Raibaud, A. & Brey, P. T. Sex ratio adjustment in *Plasmodium gallinaceum*. *Parassitologia* **41**, 153–158 (1999).
- Osgood, S. M., Eisen, R. J. & Schall, J. J. Gametocyte sex ratio of a malaria parasite: Experimental test of heritability. *J. Parasitol.* **88**, 494–498 (2002).
- Dye, C. & Godfray, H. C. F. On sex ratio and inbreeding in malaria parasite populations. *J. Theor. Biol.* **161**, 131–134 (1993).
- Nee, S., West, S. A. & Read, A. F. Inbreeding and parasite sex ratios. *Proc. R. Soc. Lond. B* **269**, 755–760 (2002).
- Read, A. F., Anwar, M., Shutler, D. & Nee, S. Sex allocation and population-structure in malaria and related parasitic protozoa. *Proc. R. Soc. Lond. B* **260**, 359–363 (1995).
- West, S. A., Smith, T. G. & Read, A. F. Sex allocation and population structure in apicomplexan (protozoa) parasites. *Proc. R. Soc. Lond. B* **267**, 257–263 (2000).
- Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, Oxford, UK, 1930).
- West, S. A., Shuker, D. M. & Sheldon, B. C. Sex-ratio adjustment when relatives interact: A test of constraints on adaptation. *Evolution Int. J. Org. Evolution* **59**, 1211–1228 (2005).
- Paul, R. E. L. *et al.* Mating patterns in malaria parasite populations of Papua New Guinea. *Science* **269**, 1709–1711 (1995).
- Conway, D. J. *et al.* High recombination rate in natural populations of *Plasmodium falciparum*. *Proc. Natl Acad. Sci. USA* **96**, 4506–4511 (1999).
- Walliker, D., Babiker, H. A. & Ranford-Cartwright, L. C. in *Malaria: Parasite Biology, Pathogenesis and Protection* (ed. Sherman, I.) 235–252 (ASM, Washington DC, 1998).
- Read, A. F. *et al.* Gametocyte sex-ratios as indirect measures of outcrossing rates in malaria. *Parasitology* **104**, 387–395 (1992).
- West, S. A., Reece, S. E. & Read, A. F. Evolution of gametocyte sex ratios in malaria and related Apicomplexan (protozoa) parasites. *Trends Parasitol.* **17**, 525–531 (2001).
- Read, A. F., Smith, T. G., Nee, S. & West, S. A. in *Sex Ratio Handbook* (ed. Hardy, I. C. W.) 314–332 (Cambridge Univ. Press, Cambridge, UK, 2002).
- Robert, V. *et al.* Malaria transmission in urban Sub-Saharan Africa. *Am. J. Trop. Med. Hyg.* **68**, 169–176 (2003).
- Robert, V. *et al.* Effect of gametocyte sex ratio on infectivity of *Plasmodium falciparum* to *Anopheles gambiae*. *Trans. R. Soc. Trop. Med. Hyg.* **90**, 621–624 (1996).
- Schall, J. J. Transmission success of the malaria parasite *Plasmodium mexicanum* into its vector: Role of gametocyte density and sex ratio. *Parasitology* **121**, 575–580 (2000).
- West, S. A., Griffin, A. S., Gardner, A. & Diggle, S. P. Social evolution theory for microorganisms. *Nature Rev. Microbiol.* **4**, 597–607 (2006).
- Gardner, A., Reece, S. E. & West, S. A. Even more extreme fertility insurance and the sex ratios of protozoan blood parasites. *J. Theor. Biol.* **223**, 515–521 (2003).
- West, S. A., Smith, T. G., Nee, S. & Read, A. F. Fertility insurance and the sex ratios of malaria and related hemosporin blood parasites. *J. Parasitol.* **88**, 258–263 (2002).
- Pickering, J., Read, A. F., Guerrero, S. & West, S. A. Sex ratio and virulence in two species of lizard malaria parasites. *Evol. Ecol. Res.* **2**, 171–184 (2000).
- Reece, S. E. & Read, A. F. Malaria sex ratios. *Trends Ecol. Evol.* **15**, 259–260 (2000).
- van Dijk, M. R. *et al.* A central role for p48/45 in malaria parasite male gamete fertility. *Cell* **104**, 153–164 (2001).
- Khan, S. M. *et al.* Proteome analysis of separated male and female gametocytes reveals novel sex-specific *Plasmodium* biology. *Cell* **121**, 675–687 (2005).
- Janse, C. J. *et al.* *In vitro* formation of ookinetes and functional maturity of *Plasmodium-berghei* gametocytes. *Parasitology* **91**, 19–29 (1985).
- Drew, D. R. & Reece, S. E. Development of reverse-transcription PCR techniques to analyse the density and sex ratio of gametocytes in genetically diverse *Plasmodium chabaudi* infections. *Mol. Biochem. Parasitol.* **156**, 199–209 (2007).
- Werren, J. H. Sex ratio adaptations to local mate competition in a parasitic wasp. *Science* **208**, 1157–1159 (1980).
- Mackinnon, M. J. & Read, A. F. Genetic relationships between parasite virulence and transmission in the rodent malaria *Plasmodium chabaudi*. *Evolution Int. J. Org. Evolution* **53**, 689–703 (1999).
- Razakandrainibe, F. G. *et al.* “Clonal” population structure of the malaria agent *Plasmodium falciparum* in high-infection regions. *Proc. Natl Acad. Sci. USA* **102**, 17388–17393 (2005).
- Nesse, R. M. & Williams, G. C. *Why We Get Sick: The New Science of Darwinian Medicine* (Times Books, New York, 1995).

44. Mehdiabadi, N. J. *et al.* Kin preference in a social microbe. *Nature* **442**, 881–882 (2006).
45. Crozier, R. H. Genetic clonal recognition abilities in marine-invertebrates must be maintained by selection for something else. *Evolution Int. J. Org. Evolution* **40**, 1100–1101 (1986).
46. Frank, S. A. A kin selection model for the evolution of virulence. *Proc. R. Soc. Lond. B* **250**, 195–197 (1992).
47. Frank, S. A. Kin selection and virulence in the evolution of protocells and parasites. *Proc. R. Soc. Lond. B* **258**, 153–161 (1994).
48. Herre, E. A. Population structure and the evolution of virulence in nematode parasites of fig wasps. *Science* **259**, 1442–1445 (1993).
49. Al-Olayan, E. M., Williams, G. T. & Hurd, H. Apoptosis in the malaria protozoan, *Plasmodium berghei*: A possible mechanism for limiting intensity of infection in the mosquito. *Int. J. Parasitol.* **32**, 1133–1143 (2002).
50. Rousset, F. & Roze, D. Constraints on the origin and maintenance of genetic kin recognition. *Evolution Int. J. Org. Evolution* **61**, 2320–2330 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. P. Waters, C. Janse and M. R. van Dijk for the genetically modified parasites, and D. H. Nussey, S. A. West, A. F. Read and A. Buckling for discussions. The Wellcome Trust, NERC, BBSRC and Royal Society provided funding.

Author Contributions S.E.R. conceived and designed the experiments, carried out the fitness consequences experiment, analysed sex ratio data and prepared the manuscript. D.R.D. developed the PCR assays, carried out the sex ratio experiments and data collection. A.G. analysed the fitness data and contributed to discussions and manuscript preparation.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.E.R. (sarah.reece@ed.ac.uk).

METHODS

The fitness consequences of sex ratio variation. To produce *P. berghei* parasites for experimental cultures, we i.p. inoculated doses of 1×10^7 parasites into mice that had been pre-treated with phenylhydrazine (60 mg kg^{-1}), to increase gametocyte production, 3 days before receiving their parasites. Parasites were harvested on day 3 PI to set up experimental cultures and also initiate the next set of infections. Each *pb48/45*-ko infection was paired with a *pb47*-ko infection and within each pair, parasites from the two lines were mixed in the appropriate ratios to produce each of our eight sex allocation treatments. Each pair of infections contributed parasites to one replicate of each of the following sex ratios (% male gametocytes): 0, 15, 30, 45, 60, 70, 90 and 100. We paired infections based on their similarity in gametocyte densities and sex ratio to keep haematocrit at 2% for all cultures. We initiated 24 pairs of infections, over 6 blocks, and chose the 19 closest matched pairs to set up our experimental cultures. By cross-factoring each pair of infections with all sex ratio treatments we avoided confounding sex ratio with infection specific factors that could influence fertilization success (for example, gametocyte age, asexual parasite density, anaemia and immune factors in serum) and could control for pair identity ($F_{18,90} = 15.39$; $P < 0.0001$). To calculate the volume of blood required from every *pb47*-ko infection (to contribute males), we calculated the density (per ml of mouse blood) of exflagellating males using a haemocytometer. To calculate the volume of blood required from every *pb48/45*-ko infection (to contribute females) we used blood smear and red blood cell density data (proportion of red blood cells containing female gametocytes \times red cells per ml). Cultures contained an average of $(0.5 \pm 0.04) \times 10^6$ gametocytes ml^{-1} and this variation was controlled for in our analysis ($F_{1,90} = 60.93$; $P < 0.0001$). We cultured 100 μl blood in RPMI with 10% calf serum, pH 8, for 20 h at 20 °C. To count ookinetes, we vortex mixed each culture before counting the number observed in a haemocytometer.

Sequencing qRT-PCR target genes in *P. chabaudi* parasite lines. We have previously developed qRT-PCR to quantify total parasites, total gametocytes and male gametocytes, based on the detection of the *P. chabaudi* common gametocyte gene 1 (*CG1*) and male gametocyte gene 1 (*MG1*). To determine what combinations of genotypes in genetically diverse infections would enable us to follow a focal genotype we sequenced *CG1* and *MG1* genes in CR, ER, DK and CW *P. chabaudi* parasite lines. A 659 bp region of the *CG1* gene and a 924 bp region of the *MG1* gene were then amplified from DNA extracted from each clone and sequenced. Our AJ specific primer sets could discriminate and quantify AJ parasites, gametocytes and male gametocytes in infections with AS and ER. Our AS specific primer sets could discriminate and quantify either AS or ER parasites, gametocytes and male gametocytes in infections with AJ but could not distinguish AS from ER parasites because AS and ER share the same *CG1* and

MG1 alleles. The other 3 genotypes each had an AS/ER allele and an AJ allele so could not be distinguished from any others, in any combination.

Experimental design. Male MF1 mice (in-house supplier, University of Edinburgh) were infected with the clonal *P. chabaudi* genotypes AS, AJ, ER, CR, CW or DK (WHO Registry of Standard Malaria Parasites, The University of Edinburgh). We initiated infections with each of our six *P. chabaudi* genotypes on their own, in combinations of two genotypes, three genotypes and all together. We infected five mice for each of 11 treatment groups as follows: (1) six groups of single-genotype infections, consisting of 1×10^6 AJ, AS, ER, CR, CW or DK parasites; (2) two groups of two-genotype infections, one group with 1×10^6 AJ + 1×10^6 AS parasites and a second group with 1×10^6 AJ + 1×10^6 ER parasites; (3) one group of three-genotype infections with 1×10^6 AJ + 1×10^6 AS + 1×10^6 ER parasites; and (4) two identical groups of six-genotype infections with 1×10^5 AJ + 1×10^5 AS + 1×10^5 ER + 1×10^5 CR + 1×10^5 CW + 1×10^5 DK parasites. All infections were sampled in the morning when the circulating parasites were in ring or early trophozoite stages from days 5 to 16 PI. Red blood cell densities were estimated using flow cytometry and reticulocyte densities were estimated from thin blood smears as previously described.

Analyses. We used R version 2.5.0 (The R Foundation for Statistical Computing; <http://www.R-project.org>) for all analyses. In all analyses sex ratios were arcsine square root transformed. Owing to the number of host deaths in mixed-genotype infections we maximized data resolution by restricting analyses of these data to days 5–12 PI. Statistics are presented from linear mixed-effects models of sex ratio variation throughout infections to overcome problems associated with pseudoreplication in longitudinal analysis. We evaluated the significance of fixed effects by comparing models using log-likelihood ratio tests following stepwise deletion of the least significant term. Specifically, we compared the change in model deviance, following term deletion, to χ^2 distributions with degrees of freedom corresponding to the difference in number of terms in the models. We simplified maximal models using maximum likelihood techniques, until only significant terms remained in the model ($\alpha < 0.05$). We then re-ran minimal models using restricted maximum likelihood to estimate the effect sizes reported in the text. We do not present results for the main effects of terms contained in significant interactions.

Models included the identity of genotypes, the number of genotypes in infections and day PI as factors (as sex ratio variation over time is nonlinear), as well as relevant interactions between these terms. Where possible we grouped genotypes together when there was no significant difference in their sex ratio patterns across infections. Covariates known or suspected to influence sex ratio were fitted, including: (1) measures of virulence and anaemia (mass, red blood cells and reticulocytes); (2) relative and absolute contributions of focal genotypes' gametocytes and asexuals to their infections; and (3) absolute number of gametocytes and asexuals in infections.

The genomic and epidemiological dynamics of human influenza A virus

Andrew Rambaut¹, Oliver G. Pybus², Martha I. Nelson³, Cecile Viboud⁴, Jeffery K. Taubenberger⁵
& Edward C. Holmes^{3,4}

The evolutionary interaction between influenza A virus and the human immune system, manifest as 'antigenic drift' of the viral haemagglutinin, is one of the best described patterns in molecular evolution. However, little is known about the genome-scale evolutionary dynamics of this pathogen. Similarly, how genomic processes relate to global influenza epidemiology, in which the A/H3N2 and A/H1N1 subtypes co-circulate, is poorly understood. Here through an analysis of 1,302 complete viral genomes sampled from temperate populations in both hemispheres, we show that the genomic evolution of influenza A virus is characterized by a complex interplay between frequent reassortment and periodic selective sweeps. The A/H3N2 and A/H1N1 subtypes exhibit different evolutionary dynamics, with diverse lineages circulating in A/H1N1, indicative of weaker antigenic drift. These results suggest a sink–source model of viral ecology in which new lineages are seeded from a persistent influenza reservoir, which we hypothesize to be located in the tropics, to sink populations in temperate regions.

Influenza is one of the most important respiratory infections of humans, responsible for 250,000 to 500,000 deaths annually¹. Of the three types of influenza virus, type A is the most virulent and is associated with seasonal (winter) epidemics in temperate regions, more persistent transmission in the tropics², and occasional large-scale global pandemics characterized by increased morbidity and mortality.

Since the global pandemic of 1918 caused by a subtype H1N1 influenza A virus, five genome segments have maintained an unbroken evolutionary history within humans—those encoding the nucleocapsid protein (NP), the matrix proteins (M1 and M2) and the nonstructural proteins (NS1 and NS2), and two encoding polymerase proteins (PB2 and PA)^{3,4}. In contrast, new haemagglutinin (HA) and neuraminidase (NA) surface glycoproteins, as well as the PB1 polymerase, have been acquired by human influenza A virus through reassortment with avian influenza viruses. These acquisitions coincided with global pandemics; the HA subtype H2 and the NA subtype N2 appeared in 1957, the HA subtype H3 emerged in 1968, and a new PB1 segment was acquired in both 1957 and 1968. Although A/H1N1 viruses re-emerged in 1977 (ref. 5) and continue to circulate, seasonal epidemics of influenza A virus since 1968 have been dominated by A/H3N2 viruses⁶, and characterized by punctuated antigenic evolution⁷.

Despite the recent availability of complete genome sequence data⁸, many aspects of the evolutionary and epidemiological dynamics of influenza A virus remain opaque. In particular, there has been no rigorous measurement of viral diversity across time, across space and among subtypes. Additionally, most studies of evolutionary dynamics have focused on single segments, without exploring the interactions among them. Although comparative studies have revealed that reassortment occurs frequently within A/H3N2 (refs 9, 10), within and between the A/H2N2 and A/H1N1 subtypes^{11–13}, and among avian influenza viruses¹⁴, the full extent of reassortment,

and how it relates to antigenic evolution, has not been determined. At the epidemiological scale, although it is apparent that A/H3N2 and A/H1N1 experience oscillating seasonal dynamics, the forces that determine this periodicity, and how they vary spatially, are unknown.

To help to resolve these issues, we examined in detail the evolutionary dynamics of influenza A virus at the genomic and epidemiological scales. Using a data set of 1,302 A/H3N2 and A/H1N1 isolates sampled over a 12-yr period and that is illustrative of large populations in temperate regions from the Northern (New York state, USA) and Southern (New Zealand) Hemispheres, we quantify the genetic diversity of influenza A virus among subtypes, genome segments and geographic locations.

Dynamics of influenza genetic diversity

The changing patterns of genetic diversity in viral isolates from New York state and New Zealand clearly reveal the seasonal dynamics of influenza A in individual temperate populations (Fig. 1). In the case of the better-sampled A/H3N2 subtype, the pattern exposed by our coalescent-based analysis is of an annual series of peaks in genetic diversity interspersed by strong genetic bottlenecks at the end of most influenza seasons. As expected, genetic diversity of the New York state epidemics peak in the Northern Hemisphere winter, whereas those of New Zealand are offset by approximately 6 months, corresponding to the Southern Hemisphere winter. (A similar pattern was observed when 123 genome sequences from Australia were added, suggesting that these patterns are common to other temperate populations, see Supplementary Information.) Furthermore, the genetic diversity of A/H3N2 is usually lower in New Zealand than in New York state, probably reflecting the smaller host population in New Zealand (most isolates were sampled from Canterbury, South Island). Differences in the size of the host population may also explain why A/H3N2 in New Zealand is sometimes less diverse than A/H1N1 in New York state, even though A/H3N2 is usually

¹Institute of Evolutionary Biology, University of Edinburgh, Ashworth Laboratories, Edinburgh EH9 3JT, UK. ²Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK. ³Center for Infectious Disease Dynamics, Department of Biology, The Pennsylvania State University, Mueller Laboratory, University Park, Pennsylvania 16802, USA. ⁴Fogarty International Center, National Institutes of Health, Bethesda, Maryland 20892, USA. ⁵Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.

epidemiologically dominant. The absolute amount of genetic diversity, even at seasonal peaks, is small compared to other rapidly evolving viruses that infect far fewer people^{15,16}, suggesting that strong natural selection, in addition to periodic bottlenecks, reduces the level of diversity that co-circulates at any time. Simulations demonstrate that our reconstructions of genetic diversity are robust to the sampling protocol (Supplementary Information).

More notable is that, in both populations, A/H1N1 seasons exhibiting highly defined peaks in diversity typically coincide with weakly defined peaks in A/H3N2 diversity (that is, the measures of the epidemic 'peakedness' for A/H3N2 and A/H1N1 are negatively correlated; Wilcoxon signed-rank test, $W = 348$, $n = 32$, $P < 0.002$; Supplementary Figs 1 and 2). This implies an evolutionary interaction between subtypes; for example, that A/H1N1 epidemics are suppressed by herd immunity when A/H3N2 is dominant, or that A/H3N2 out-competes A/H1N1, perhaps owing to greater replicative fitness. Consistent with our observation, A/H1N1 only dominates in seasons following unusually mild H3N2 epidemics¹⁷, and infection with one subtype can protect against reinfection with the other in sequential epidemics¹⁸. Furthermore, since 1977, A/H1N1 epidemics exhibit lower mortality rates than A/H3N2 epidemics (refs 19 and 20) and are less spatially synchronized^{21–23}. In both hemispheres, A/H1N1 seems less prone to the seasonal genetic bottlenecks that characterize A/H3N2, suggesting that genetically diverse A/H1N1 lineages are better able to coexist. This might indicate that antigenic selection acts with less potency on A/H1N1, manifest as lower rates of amino acid fixation in HA1 (ref. 17), so that the selective turnover of lineages occurs more slowly^{17,24}.

The persistence of viral diversity between epidemic peaks for A/H3N2 and A/H1N1 has two explanations: that chains of infection are surviving in each population and across inter-epidemic intervals, or that genetic diversity is imported into temperate populations each year. Phylogenetic evidence strongly weighs against the former, because there are few direct phylogenetic links between influenza A viruses sampled in successive seasons from specific locations, as expected if *in situ* evolution was commonplace^{25,26}. Furthermore, a high risk of stochastic extinction makes the repeated survival of small chains of infection between successive epidemics unlikely. Additionally, the dates of segment ancestry in the New York state 1999–2000 season reappeared precisely in the 2002–2003 season, after a major bottleneck in the A/H1N1-dominant 2000–2001 season, strongly suggesting that diversity was maintained in a reservoir population.

Extrapolating the evolutionary dynamics of influenza in New York state and New Zealand to other comparable populations in the Northern and Southern Hemispheres leads us to propose a 'source-sink' model for the global ecology of influenza A virus

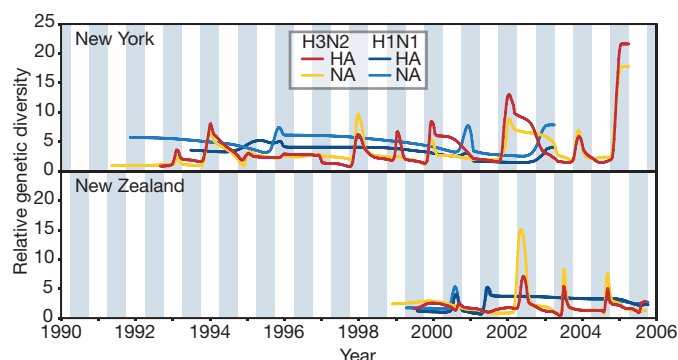


Figure 1 | Population dynamics of genetic diversity in influenza A virus. Bayesian skyline plots of the HA and NA segments for the A/H3N2 and A/H1N1 subtypes in New York state (top) and New Zealand (bottom). The horizontal shaded blocks represent the winter seasons. The y-axes represent a measure of relative genetic diversity (see Methods for details). The shorter timescale of New Zealand skyline plot is due to the shorter sampling period.

(Fig. 2). Continual—but largely unidirectional—gene flow from a common source population (or a linked network of source populations^{24,27}) provides the viruses that ignite each epidemic in populations of the Northern and Southern Hemispheres. Although southern China has been proposed as the epicentre of influenza A virus²⁸, it is possible that tropical regions generally represent ideal source populations because of extended viral transmission². A necessary consequence of this model is that selection-driven antigenic drift will be much more efficient in the source population, the long-term effective size of which is maintained by a high background infection rate and by the absence of the severe population bottlenecks associated with temperate localities². Hence, the observation of antigenic change in temperate populations is a secondary effect of selection within, and gene flow from, the source population. This explains the paradox of there being frequent positive selection on HA across seasons at a global scale^{29–31} but little evidence for antigenic drift at the scale of individual seasonal epidemics in those temperate populations studied so far^{26,32}. To test this hypothesis, it will be essential to obtain more influenza virus samples from tropical regions, such as South-East Asia.

Genomic diversity and reassortment

To determine the genome-wide evolutionary dynamics of human influenza A, we inferred the population genetic history of each segment of A/H3N2 viruses sampled from New York state (the largest data set). This analysis demonstrated that, at any given time, individual segments can differ substantially in their relative genetic diversity and hence in phylogenetic history (Supplementary Fig. 3).

To determine the causes of these differences in genetic diversity, we estimated the time to the most recent common ancestor (TMRCA) of each segment for each influenza season (Fig. 3a; the highest posterior density intervals for each estimate are shown in Supplementary Fig. 4). Most TMRCAs fall well before the start of the season from which they were sampled, such that multiple lineages persist across multiple epidemic troughs, consistent with our source-sink model. TMRCAs also vary among segments and among years, reflecting the interacting processes of genomic reassortment, natural selection and gene flow. For example, the TMRCAs of the NP segment are usually greater than those of the HA segment, indicating that genetic diversity persists for longer in the former. These changing patterns of diversity provide a unique insight into genome-wide evolutionary processes. In particular, the 2001–2002 season is characterized by reduced TMRCAs across the entire A/H3N2 genome (following a dominant A/H1N1 season), such that all segments have approximately the same

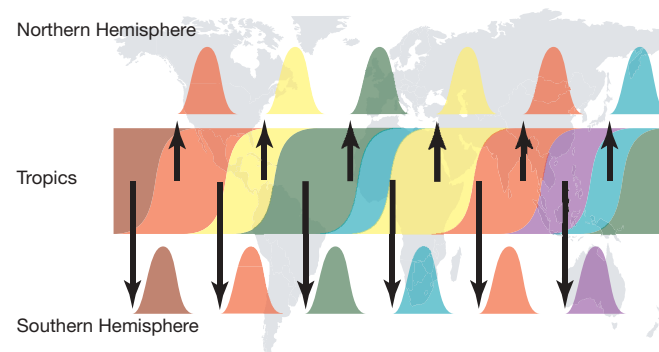


Figure 2 | A 'source-sink' model for the evolution of influenza A virus. Viral genetic and antigenic diversity (shown by different colours) is continuously generated in a reservoir, or 'source' population, perhaps represented by the tropics, before being exported to 'sink' populations in the Northern and Southern Hemispheres as shown by the arrows. The continuous transmission of influenza A virus in the source population, and hence its larger effective population size, allows natural selection for antigenic diversity to proceed more efficiently than in the sink populations that are afflicted by major seasonal bottlenecks.

TMRCAs, strongly suggesting a genome-wide selective sweep. In the following 2002–2003 season, genetic diversity reappears at the genomic scale, so that all segments have TMRCAs dating no later than 1999–2000. Subsequently, in 2003–2004, the HA segment alone has a reduced TMRCAs, indicating that it underwent a selective sweep, with reassortment maintaining the diversity of other segments. Another genome-wide selective sweep is reflected in the tight TMRCAs of 2004–2005.

An analysis of the HA and NA from A/H1N1 in New York state and New Zealand reveals a similar pattern of temporal fluctuations in diversity (Fig. 3b). However, the TMRCAs are often much greater than those observed for A/H3N2 in the same year, providing further evidence that greater variation persists among epidemic seasons in A/H1N1 than in A/H3N2. This may be due to a smaller effective population size or weaker immune selection in A/H1N1, such that natural selection is less effective at fixing advantageous variants.

The occurrence of reassortment and natural selection also coincides with changes in HA antigenicity (Fig. 3a). The earliest antigenic transition in our data is between Beijing/1993 (BE93)-like and

Wuhan/1995 (WU95)-like viruses⁷. This is associated with a sharp increase in the TMRCAs of PB2 in the 1996–1997 season, owing to the reappearance of a lineage present until 1994–1995, coupled with reassortment. Similarly, the WU95 to Sydney/1997-like virus (SY97) transition coincides with a marked reduction in the genetic diversity of all segments other than M1/2, in which reassortment has unlinked this segment from the rest of the genome, whereas a new HA lineage is acquired at the time of the SY97 to Fujian/2002-like virus (FU02) antigenic change³³. Finally, the genome-wide selective sweep of 2004–2005 coincides with the FU02 to California/2004-like virus (CA04) antigenic transition.

Evolutionary interactions among segments

We used multivariate statistics to summarize the differences in phylogenetic history among segments for A/H3N2 viruses from New York state (Fig. 4 and Supplementary Fig. 12). Each cloud of points of equal colour corresponds to a specific segment of the influenza virus genome, with the spread of points within a cloud representing the statistical uncertainty in the phylogenetic history of that segment. Clearly, the differences in history among segments are greater than the estimation error for each segment. The most prominent result is the divergent position of HA, particularly in relation to NA, even though both encode surface glycoproteins. These differences in evolutionary history are compatible with frequent reassortment, suggesting that the HA segment is continually placed in new genomic backgrounds, perhaps greatly affecting virus fitness.

In contrast, there is clear similarity in the phylogenetic histories of the HA and M1/2 segments, compatible with the observations that HA interacts structurally with M1 in viral assembly and that the M2 ion channel balances the pH for optimal HA fusion^{34–38}. There is also a similarity in phylogenetic history among the NS and NP segments, suggesting that they are subject to varying degrees of physical linkage. Elucidating the physiological basis for these linkage patterns will undoubtedly provide information about the emergence of pandemic strains of influenza virus.

Further insights into segment interactions are provided by an analysis of evolutionary rates in A/H3N2 isolates from New York state

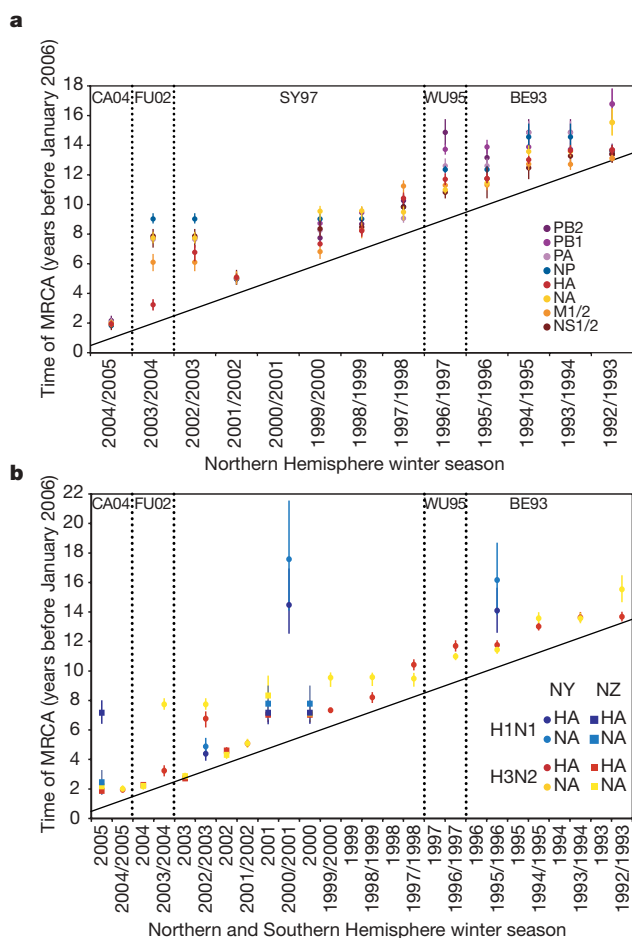


Figure 3 | Population genetic history of human influenza A virus. **a, b,** Time to the most recent common ancestor (TMRCAs) of each genomic segment for A/H3N2 isolates circulating each season in New York state (**a**) and HA and NA genomic segments for the A/H3N2 and A/H1N1 isolates circulating in New York state and New Zealand (**b**). The values shown represent the mean and 95% highest posterior density intervals for TMRCAs estimated across the trees sampled using bayesian MCMC analyses. Explicit posterior probabilities of whether any season has a TMRCAs that is older than that of each preceding season are given in Supplementary Table 1. The diagonal line goes through 1st January of each season (1st July for New Zealand), approximating the seasonal mid-point. The timescale of major changes in antigenic (HA) type in the United States is also depicted. In the Northern Hemisphere, individual influenza seasons straddle two years, whereas in the Southern Hemisphere they are contained within one calendar year.

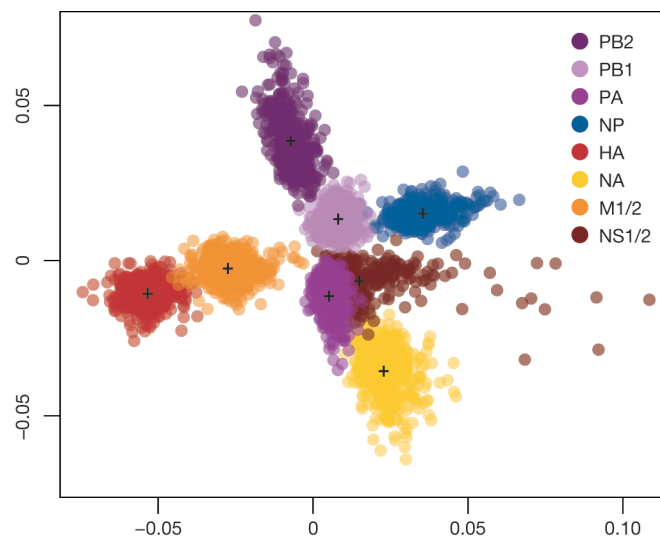


Figure 4 | Differences in phylogenetic history among segments of H3N2 influenza A virus. A multi-dimensional scaling plot of distances between samples of 500 trees for each segment of A/H3N2 viruses sampled from New York state. The tree-to-tree distances are calculated as one minus the correlation coefficient of the vector of TMRCAs for each season. Each cloud of points of equal colour represents the statistical uncertainty in the phylogenetic history of an individual genomic segment (with the centroids of each segment distribution indicated). Considering each comparison in a pairwise manner between segments reveals a similar pattern (Supplementary Fig. 12).

(Supplementary Fig. 5). We identified a major rate difference between those segments expressed on the surface of the virion (high rates) and those that only have internal functions (low rates); the exception to this was the NS segment, for which only the NS2 protein is a minor virion component. As expected given its role in immune evasion, the HA protein exhibits the highest overall rate of evolutionary change (mean of 5.72×10^{-3} nucleotide substitutions per site per year (subs per site per year), 95% highest posterior density interval $5.17, 6.28 \times 10^{-3}$ subs per site per year), as well as the highest rate of nonsynonymous substitution, indicated by the relative rates at the first and second versus the third codon positions. More surprisingly, the M1/2, NS1/2 and NA segments evolve at rates overlapping those of HA (mean of 5.20×10^{-3} subs per site per year (95% highest posterior density interval $4.36, 6.07 \times 10^{-3}$ subs per site per year), 5.39×10^{-3} ($4.48, 6.35 \times 10^{-3}$) and 5.41×10^{-3} ($4.81, 5.99 \times 10^{-3}$), respectively), although with lower rates of nonsynonymous change. In contrast, the three segments that constitute the viral polymerase, as well as the RNA-binding NP, exhibit markedly lower rates at all sites and at the first and second codon positions only, revealing stronger selective constraints against amino acid variation.

Because the NA and HA segments have different phylogenetic histories, their similarities in overall evolutionary rate are unlikely to be wholly due to genetic linkage. Hence, the NA is also likely to be subject to relatively strong positive selection, either for antigenic variation or for functional compatibility with HA³⁹. Immune-mediated selection may also explain the increased rates in M1/2 and NS1/2. The extracellular domain of the M2 ion channel is a key target in the development of a 'universal' influenza vaccine⁴⁰, whereas the M1 protein is the major component of the viral capsid and a potential target of cellular immune responses⁴¹. Likewise, NS1 encodes a pleiotropic, nonstructural protein that downregulates double-stranded-RNA-induced antiviral responses⁴², or that perhaps co-evolves with the PB1 segment as observed in the last two pandemics^{30,43}.

Conclusions

Our analyses suggest that the evolutionary dynamics of influenza A virus are shaped by a complex interplay between rapid mutation, frequent reassortment, widespread gene flow, natural selection (occasionally generating genome-wide selective sweeps), functional interactions among segments, and global epidemiological dynamics. Although the increased rates of evolution in HA support models in which this protein is subject to strong selection for immune escape, the high rates of evolutionary change in other virion-associated proteins (that have different phylogenetic histories to HA) indicate that they are likewise subject to strong positive selection. It is possible that a significant component of this adaptation involves optimizing the functional compatibility of segments; for example, a new HA variant may not increase fitness unless it is linked to functionally compatible segments. Such a phenomenon may explain why the lineage leading to the FU02 antigenic type did not dominate the viral population until a few years after its initial appearance and co-incident with an HA reassortment event³³.

Equally notable was the high degree of genomic reassortment and its association with periodic selective sweeps. It is therefore crucial to consider the process of antigenic drift within the context of frequent reassortment and genome-wide epistatic interactions. Reassortment places antigenically novel HA variants into different genomic backgrounds, a fraction of which may restore, or even increase, the viral replicative fitness that may have been lost as a result of the change in HA. Although it is unlikely that the reassortment rate *per se* varies among segments, it is probable that most reassortments involving certain segment combinations are deleterious—such as those involving HA and M1/2, which seem to be relatively tightly linked—and are removed by purifying selection, giving rise to the similarities and differences in phylogenetic history discovered here.

In addition to genome-wide interactions, it is essential to consider the complex spatial epidemiological dynamics of influenza if we are to fully understand antigenic evolution. We observed consistent dynamical patterns in two populations illustrative of temperate regions in the Northern and Southern Hemispheres, together with the persistence of viral lineages across multiple epidemics. To resolve these apparently contradictory observations, we propose the existence of a continuous reservoir or source population, within which the strong selection for antigenic change takes place. Such complexity necessarily means that the long-term success of any individual lineage of influenza virus is dependent not only on its antigenic properties but also on its replicative capacity, its transmissibility and the environmental factors that perhaps underlie the seasonality of influenza in temperate regions. In this context, it is important to determine the precise reasons why antigenic evolution proceeds at a reduced pace in A/H1N1, such that multiple lineages co-circulate globally, even though this subtype is occasionally the major cause of seasonal influenza.

METHODS SUMMARY

Complete genome sequences of A/H3N2 and A/H1N1 influenza viruses from New York state, USA, and New Zealand, were collated from the NCBI Influenza Database as part of the Influenza Genome Sequencing Project^{8,44}. Sequence alignments representing each genomic segment, viral subtype and sampling location were assembled. Evolutionary dynamics were estimated for each data set using an established bayesian Markov chain Monte Carlo (MCMC) approach^{45,46} that incorporates the exact day of viral sampling. Each MCMC analysis resulted in marginal posterior estimates of: the rate of nucleotide substitution (Supplementary Fig. 5); the relative evolutionary rate at the first and second versus the third codon positions (Supplementary Fig. 5); and the dynamics of population genetic diversity through time (see Fig. 1 and Supplementary Fig. 4). Additionally, an estimate of the posterior distribution of genealogies that relate the sequence data was inferred concurrently for each alignment. From these distributions, we obtained the marginal posterior estimate of the TMRCA of each genomic segment in each influenza season (Fig. 3a, b). Differences between the node-height distributions of each posterior set of genealogies were visualized using multidimensional scaling (Fig. 4). The node heights and topological structure of each genealogical distribution was summarized using the maximum clade support approach (Supplementary Fig. 3).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 3 July 2007; accepted 19 March 2008.

Published online 16 April 2008.

1. WHO. Fact sheet Number 211. Influenza. (<http://www.who.int/mediacentre/factsheets/fs211/>) (2003).
2. Viboud, C., Alonso, W. J. & Simonsen, L. Influenza in tropical regions. *PLoS Med.* **3**, e89 (2006).
3. Reid, A. H., Taubenberger, J. K. & Fanning, T. G. Evidence of an absence: the genetic origins of the 1918 pandemic influenza virus. *Nature Rev. Microbiol.* **2**, 909–914 (2004).
4. Taubenberger, J. K. *et al.* Characterization of the 1918 influenza virus polymerase genes. *Nature* **437**, 889–893 (2005).
5. Scholtissek, C., von Hoyningen, V. & Rott, R. Genetic relatedness between the new 1977 epidemic strains (H1N1) of influenza and human influenza strains isolated between 1947 and 1957 (H1N1). *Virology* **89**, 613–617 (1978).
6. Taubenberger, J. K. & Morens, D. M. 1918 Influenza: the mother of all pandemics. *Emerg. Infect. Dis.* **12**, 15–22 (2006).
7. Smith, D. J. *et al.* Mapping the antigenic and genetic evolution of influenza virus. *Science* **305**, 371–376 (2004).
8. Ghedin, E. *et al.* Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. *Nature* **437**, 1162–1166 (2005).
9. Lindstrom, S. E. *et al.* Phylogenetic analysis of the entire genome of influenza A (H3N2) viruses from Japan: evidence for genetic reassortment of the six internal genes. *J. Virol.* **72**, 8021–8031 (1998).
10. Schweiger, B., Bruns, L. & Meixenberger, K. Reassortment between human A(H3N2) viruses is an important evolutionary mechanism. *Vaccine* **24**, 6683–6690 (2006).
11. Gregory, V. *et al.* Emergence of influenza A H1N2 reassortant viruses in the human population during 2001. *Virology* **300**, 1–7 (2002).
12. Lindstrom, S. E., Cox, N. J. & Klimov, A. Genetic analysis of human H2N2 and early H3N2 influenza viruses, 1957–1972: evidence for genetic divergence and multiple reassortment events. *Virology* **328**, 101–119 (2004).

13. Young, J. F. & Palese, P. Evolution of human influenza A viruses in nature: recombination contributes to genetic variation of H1N1 strains. *Proc. Natl Acad. Sci. USA* **76**, 6547–6551 (1979).
14. Macken, C. A., Webby, R. J. & Bruno, W. J. Genotype turnover by reassortment of replication complex genes from avian influenza A virus. *J. Gen. Virol.* **87**, 2803–2815 (2006).
15. Carrington, C., Foster, J., Pybus, O., Bennett, S. & Holmes, E. Invasion and maintenance of dengue virus type 2 and type 4 in the Americas. *J. Virol.* **79**, 14680–14687 (2005).
16. Lemey, P., Rambaut, A. & Pybus, O. HIV evolutionary dynamics within and among hosts. *AIDS Rev.* **8**, 125–140 (2006).
17. Wolf, Y. I., Viboud, C., Holmes, E. C., Koonin, E. V. & Lipman, D. J. Long intervals of stasis punctuated by bursts of positive selection in the seasonal evolution of influenza A virus. *Biol. Direct.* **1**, 34 (2006).
18. Sonoguchi, T., Naito, H., Hara, M., Takeuchi, Y. & Fukumi, H. Cross-subtype protection in humans during sequential, overlapping, and/or concurrent epidemics caused by H3N2 and H1N1 influenza viruses. *J. Infect. Dis.* **151**, 81–88 (1985).
19. Kaji, M., Watanabe, A. & Aizawa, H. Differences in clinical features between influenza A H1N1, A H3N2, and B in adult patients. *Respirology* **8**, 231–233 (2003).
20. Wright, P. F., Thompson, J. & Karzon, D. T. Differing virulence of H1N1 and H3N2 influenza strains. *Am. J. Epidemiol.* **112**, 814–819 (1980).
21. Hay, A. J., Gregory, V., Douglas, A. R. & Lin, Y. P. The evolution of human influenza viruses. *Phil. Trans. R. Soc. Lond. B* **356**, 1861–1870 (2001).
22. Finkelman, B. *et al.* Global patterns in seasonal activity of influenza A/H3N2, A/H1N1, and B from 1997 to 2005: viral coexistence and latitudinal gradients. *PLoS ONE* **2**, e1296 (2007).
23. Viboud, C. *et al.* Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* **312**, 447–451 (2006).
24. Ferguson, N. M., Galvani, A. P. & Bush, R. M. Ecological and immunological determinants of influenza evolution. *Nature* **422**, 428–433 (2003).
25. Nelson, M. I., Simonsen, L., Viboud, C., Miller, M. A. & Holmes, E. C. Phylogenetic analysis reveals the global circulation pattern of seasonal influenza A viruses. *PLoS Pathog.* **3**, e131 (2007).
26. Nelson, M. I. *et al.* Stochastic processes are key determinants of short-term evolution in influenza A virus. *PLoS Pathog.* **2**, e125 (2006).
27. Alonso, W. J. *et al.* Seasonality of influenza in Brazil: a traveling wave from the Amazon to the subtropics. *Am. J. Epidemiol.* **165**, 1434–1442 (2007).
28. Shortridge, K. F. Is China an influenza epicentre? *Chin. Med. J. (Engl.)* **110**, 637–641 (1997).
29. Bush, R. M., Fitch, W. M., Bender, C. A. & Cox, N. J. Positive selection on the H3 hemagglutinin gene of human influenza virus A. *Mol. Biol. Evol.* **16**, 1457–1465 (1999).
30. Fitch, W. M., Leiter, J. M., Li, X. Q. & Palese, P. Positive Darwinian evolution in human influenza A viruses. *Proc. Natl Acad. Sci. USA* **88**, 4270–4274 (1991).
31. Shih, A. C., Hsiao, T. C., Ho, M. S. & Li, W. H. Simultaneous amino acid substitutions at antigenic sites drive influenza A hemagglutinin evolution. *Proc. Natl Acad. Sci. USA* **104**, 6283–6288 (2007).
32. Lavenau, A. *et al.* Detailed analysis of the genetic evolution of influenza virus during the course of an epidemic. *Epidemiol. Infect.* **134**, 514–520 (2006).
33. Holmes, E. C. *et al.* Whole-genome analysis of human influenza A virus reveals multiple persistent lineages and reassortment among recent H3N2 viruses. *PLoS Biol.* **3**, e300 (2005).
34. Ali, A., Avalos, R. T., Ponimaskin, E. & Nayak, D. P. Influenza virus assembly: effect of influenza virus glycoproteins on the membrane association of M1 protein. *J. Virol.* **74**, 8709–8719 (2000).
35. Barman, S., Ali, A., Hui, E. K., Adhikary, L. & Nayak, D. P. Transport of viral proteins to the apical membranes and interaction of matrix protein with glycoproteins in the assembly of influenza viruses. *Virus Res.* **77**, 61–69 (2001).
36. Enami, M. & Enami, K. Influenza virus hemagglutinin and neuraminidase glycoproteins stimulate the membrane association of the matrix protein. *J. Virol.* **70**, 6653–6657 (1996).
37. Ohuchi, M. *et al.* Rescue of vector-expressed fowl plague virus hemagglutinin in biologically active form by acidotropic agents and coexpressed M2 protein. *J. Virol.* **68**, 920–926 (1994).
38. Takeuchi, K. & Lamb, R. A. Influenza virus M2 protein ion channel activity stabilizes the native form of fowl plague virus hemagglutinin during intracellular transport. *J. Virol.* **68**, 911–919 (1994).
39. Mitnaul, L. J. *et al.* Balanced hemagglutinin and neuraminidase activities are critical for efficient replication of influenza A virus. *J. Virol.* **74**, 6015–6020 (2000).
40. Gerhard, W., Mozdzanowska, K. & Zharikova, D. Prospects for universal influenza virus vaccine. *Emerg. Infect. Dis.* **12**, 569–574 (2006).
41. Berkhoff, E. G. *et al.* Functional constraints of influenza A virus epitopes limit escape from cytotoxic T lymphocytes. *J. Virol.* **79**, 11239–11246 (2005).
42. Krug, R. M., Yuan, W., Noah, D. L. & Latham, A. G. Intracellular warfare between human influenza viruses and human cells: the roles of the viral NS1 protein. *Virology* **309**, 181–189 (2003).
43. Ludwig, S., Schultz, U., Mandler, J., Fitch, W. M. & Scholtissek, C. Phylogenetic relationship of the nonstructural (NS) genes of influenza A viruses. *Virology* **183**, 566–577 (1991).
44. *Influenza Genome Sequencing Project* (<http://www.niaid.nih.gov/dmid/genomes/mcs/influenza.htm>) (2007).
45. Drummond, A. J., Nicholls, G. K., Rodrigo, A. G. & Solomon, W. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* **161**, 1307–1320 (2002).
46. Drummond, A. J. & Rambaut, A. BEAST: bayesian evolutionary analysis sampling trees. *BMC Evol. Biol.* **7**, 214 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This research was supported in part by the Intramural Research Program of the NIH, Fogarty International Center, the National Institute of Allergy and Infectious Diseases and the National Institute of General Medical Sciences. A.R. and O.G.P. are supported by The Royal Society of London. A.R. works as a part of the Interdisciplinary Centre for Human and Avian Influenza Research (ICHAIR).

Author Contributions A.R. and E.C.H. conceived the study. A.R. and O.G.P. undertook the genome sequence analyses and development of methods. E.C.H., A.R. and O.G.P. wrote the paper. M.I.N., C.V. and J.K.T. provided epidemiological and virological data. All authors interpreted the results and commented on the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.R. (a.rambaut@ed.ac.uk) and E.C.H. (ech15@psu.edu).

METHODS

To examine the interaction between the H1N1 and H3N2 subtypes of human influenza A virus and the seasonal fluctuations in the Northern and Southern Hemispheres, we compiled the HA and NA genes (coding regions only) for A/H1N1 in New York state (81 isolates sampled between 1995 and 2003) and in New Zealand (127 isolates sampled between 2000 and 2005), as well as for A/H3N2 in New York state (687 isolates sampled between 1993 and 2005) and in New Zealand (407 isolates between 2000 and 2005). Influenza is strongly seasonal in both localities, with few reported cases outside of their respective winters (data available at the World Health Organization 'FluNet' surveillance system; <http://gamapserver.who.int/GlobalAtlas/home.asp>). To examine genome-wide evolutionary processes in more detail, we conducted an equivalent analysis on each segment of the 687 isolates of A/H3N2 sampled from New York state. For each segment, the protein-coding regions were extracted and aligned. For segments with more than one gene (M1/2 and NS1/2), both genes were concatenated, with overlapping codons included only once (see Supplementary Information). In all cases, alignment was unambiguous, resulting in the following data sets: PB2 (2,277 nucleotides), PB1 (2,271 nucleotides), PA (2,148 nucleotides), HA (1,698 nucleotides), NP (1,494 nucleotides), NA (1,407 nucleotides), M1/2 (979 nucleotides) and NS1/2 (835 nucleotides). A full list of viral isolates, along with their GenBank accession numbers, is available as Supplementary Information.

We used a flexible demographic model to examine the seasonal changes in genetic diversity over the sampling time span⁴⁷. This approach, based on the coalescent⁴⁸, allows the overall genetic diversity of a population to be estimated from a small sample of genetic sequences. An estimate of the change in diversity over time can be obtained for a given genealogy⁴⁷. Individual estimates obtained from many different genealogies are combined, with genealogies and parameter values sampled according to their relative probability, given the sequence data^{45,46}. In the absence of natural selection, the genetic diversity measure obtained reflects the change in effective number of infections over time ($N_e\tau$, where τ is the average generation time). Because strong natural selection has previously been demonstrated for influenza A virus, we interpret these plots as measures of relative genetic diversity. The MCMC approach implemented in the BEAST package has the advantage that it is not conditioned on a single, and potentially unrepresentative, estimate of the underlying genealogy. It provides marginal posterior estimates of all the parameters of the substitution process including the overall rate of molecular evolution and the times of all the nodes in

the genealogy—a timescale that is informed by the known dates (days) of sampling of the sequences and the amount of genetic change between them^{45,49}. Indeed, the use of sequences that have been sampled over many seasons allows the reconstruction of fluctuating dynamics limited by the date of the earliest sample (see Supplementary Information).

For these analyses, we used a nucleotide-substitution model that accommodates the different rates of and constraints on evolution at the different codon positions⁵⁰. This allows us to estimate the overall rate of molecular evolution for each gene and the relative rate of first and second codon position to the third—a measure of the constraint against amino acid change that correlates strongly with the ratio of nonsynonymous to synonymous nucleotide substitutions⁵⁰.

We investigated the effect of reassortment among segments by comparing the correlation of phylogenetic trees sampled independently for each segment. For each tree, we identified the most recent common ancestor of the samples from each season and plotted a marginal posterior probability distribution for the time of these nodes (Fig. 3). Furthermore, we estimated one minus the correlation coefficient of these TMRCAs both within and between a sample of 500 trees from the posterior distribution obtained from the MCMC analysis of each genomic segment. These relationships were plotted using the multi-dimensional scaling dimension reduction technique (Fig. 4) and pair-wise in Supplementary Fig. 12.

For each segment alignment, we used two independent runs of BEAST of 100,000,000 steps sampling trees and parameters every 10,000 steps. The two runs were compared to confirm that both had converged and were sampling the same distributions, and were then combined (having removed 10% as 'burn-in'). All the results here were summarized from the remaining 18,000 sampled trees or parameter values.

47. Drummond, A. J., Rambaut, A., Shapiro, B. & Pybus, O. G. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* **22**, 1185–1192 (2005).
48. Kingman, J. The coalescent. *Stochastic Process. Appl.* **13**, 235–248 (1982).
49. Rambaut, A. Estimating the rate of molecular evolution: Incorporating non-contemporaneous sequences into maximum likelihood phylogenies. *Bioinformatics* **16**, 395–399 (2000).
50. Shapiro, B., Rambaut, A. & Drummond, A. J. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol. Biol. Evol.* **23**, 7–9 (2006).

ARTICLES

A microbial symbiosis factor prevents intestinal inflammatory disease

Sarkis K. Mazmanian^{1*}, June L. Round^{1*} & Dennis L. Kasper^{2,3}

Humans are colonized by multitudes of commensal organisms representing members of five of the six kingdoms of life; however, our gastrointestinal tract provides residence to both beneficial and potentially pathogenic microorganisms. Imbalances in the composition of the bacterial microbiota, known as dysbiosis, are postulated to be a major factor in human disorders such as inflammatory bowel disease. We report here that the prominent human symbiont *Bacteroides fragilis* protects animals from experimental colitis induced by *Helicobacter hepaticus*, a commensal bacterium with pathogenic potential. This beneficial activity requires a single microbial molecule (polysaccharide A, PSA). In animals harbouring *B. fragilis* not expressing PSA, *H. hepaticus* colonization leads to disease and pro-inflammatory cytokine production in colonic tissues. Purified PSA administered to animals is required to suppress pro-inflammatory interleukin-17 production by intestinal immune cells and also inhibits *in vitro* reactions in cell cultures. Furthermore, PSA protects from inflammatory disease through a functional requirement for interleukin-10-producing CD4⁺ T cells. These results show that molecules of the bacterial microbiota can mediate the critical balance between health and disease. Harnessing the immunomodulatory capacity of symbiosis factors such as PSA might potentially provide therapeutics for human inflammatory disorders on the basis of entirely novel biological principles.

Many human disorders seem to require a critical—and often unknown—environmental component. Inflammatory bowel diseases (IBDs) such as Crohn's disease and ulcerative colitis represent aberrant immune responses of the human gastrointestinal tract with adverse clinical outcomes¹. Abundant clinical and laboratory research has shown that, in IBD, commensal bacteria harboured within mammalian intestines are the targets of inflammatory responses^{1–4}. Antibiotic treatment alleviates intestinal inflammation in humans and experimental animals⁵. Germ-free re-derivation of animals genetically susceptible to colitis prevents development of intestinal inflammation⁶. Transfer of CD4⁺ T-cell clones specific for bacterial antigens induces disease in recipient animals⁷. In humans and other animals, inflammatory responses are apparently directed towards specific subsets of commensal organisms that have pathogenic potential but are not typically infectious pathogenic agents. All mammals harbour these species; why inflammation ensues only in those affected by IBD is unknown. Some investigators have predicted that, in addition to genetic factors, an imbalance in the normal microbiota without acquisition of an infectious organism is at least partially responsible for IBD⁸. Metagenomic studies have shown that entire classes of bacteria are lost or over-represented as part of the IBD process⁹. Perhaps in certain disorders where environmental factors are implicated, an imbalance between commensal bacteria with pathogenic potential (which we term pathobionts) and symbionts (commensal bacteria with beneficial potential) in the microbiota has a role in pathogenesis.

Humans maintain a lifelong association with innumerable commensal microorganisms that inhabit almost every environmentally exposed surface of the body. The gastrointestinal tract harbours > 10¹⁴ microorganisms of ~1,000 species¹⁰. Collectively, the intestinal microbiota represents a 'forgotten organ' that can execute many physiological functions and thus may profoundly influence human

biology. Germ-free animals, born and raised under sterile conditions, have defects in the development of intestinal tissues, show reduced vascular, nutritional and endocrine function, and are more susceptible to infection than conventionally colonized animals^{11,12}. Both gastrointestinal and systemic immune responses are deficient in the absence of commensal microorganisms¹³. Thus, mammals seem to depend on the microbiota to promote development and differentiation of host tissue¹⁴. Because of the complexity of the interactions of this diverse consortium of microorganisms with the mammalian host, the molecules responsible for host–microbe communication remain almost entirely unknown. As the microbiota has been implicated in disease, an understanding of these molecules may benefit human health¹⁵.

We have demonstrated that germ-free animals have defects in CD4⁺ T-cell development and that the human commensal bacterium *Bacteroides fragilis* corrects these deficiencies through the expression of PSA¹³. The precise immune-cell subset affected by PSA has not yet been identified. CD4⁺ T cells of the mammalian immune system can be generally divided into a naive ('uneducated') CD4⁺CD45Rb^{high} population and an antigen-experienced ('educated') CD4⁺CD45Rb^{low} population¹⁶. We found that splenic cells from germ-free animals included a smaller proportion of CD4⁺CD45Rb^{low} T cells than those from age-matched conventional mice with a complete bacterial microbiota (Fig. 1a). We examined the ability of *B. fragilis* to correct deficiencies in the CD4⁺CD45Rb^{low} T-cell population. Mono-colonization of germ-free mice with wild-type *B. fragilis* alone restores the CD4⁺CD45Rb profile to that found in animals with a complete bacterial microbiota (Fig. 1a; left panels). Notably, colonization with a mutant strain defective in the ability to produce PSA (*B. fragilis* ΔPSA) did not generate an expansion of the CD4⁺CD45Rb^{low} T-cell population (Fig. 1a; bottom right panel). It is well established that the latter population possesses potent

¹Division of Biology, California Institute of Technology, Pasadena, California 91125, USA. ²Channing Laboratory, Brigham & Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. ³Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA.

*These authors contributed equally to this work.

anti-inflammatory properties and confers protection in animal models of inflammation¹⁷. These results suggest that *B. fragilis* may have evolved a molecular strategy to mediate protection from inflammation during host–bacterial mutualism.

Protection from colitis by *B. fragilis*

We used the well-established CD4⁺CD45Rb transfer model of experimental colitis¹⁸ to investigate whether *B. fragilis* colonization protects animals from inflammatory disease. In this model, pathogenic CD4⁺CD45Rb^{high} T cells are separated from protective CD4⁺CD45Rb^{low} cells and transferred into specific pathogen-free *Rag*^{−/−} mice. On cell transfer, mice were colonized with *Helicobacter hepaticus*^{8,19}, a pathobiont that is a benign commensal in wild-type animals but an opportunistic pathogen causing colitis in immunocompromised mice. After 8 weeks, animals were killed and colitis was assessed using a standard scoring system²⁰. The pathology scores show that *H. hepaticus* colonization and CD4⁺CD45Rb^{high} T-cell transfer are sufficient to induce severe colitis in *Rag*^{−/−} mice (Fig. 1b, first column), as previously reported^{19,21}. Co-colonization with wild-type *B. fragilis* resulted in significant protection from disease (second column); conversely, co-colonization with *B. fragilis* ΔPSA does not offer protection (third column).

Tissue damage in colitis is widely believed to result from production of inflammatory cytokines in response to commensal bacteria²².

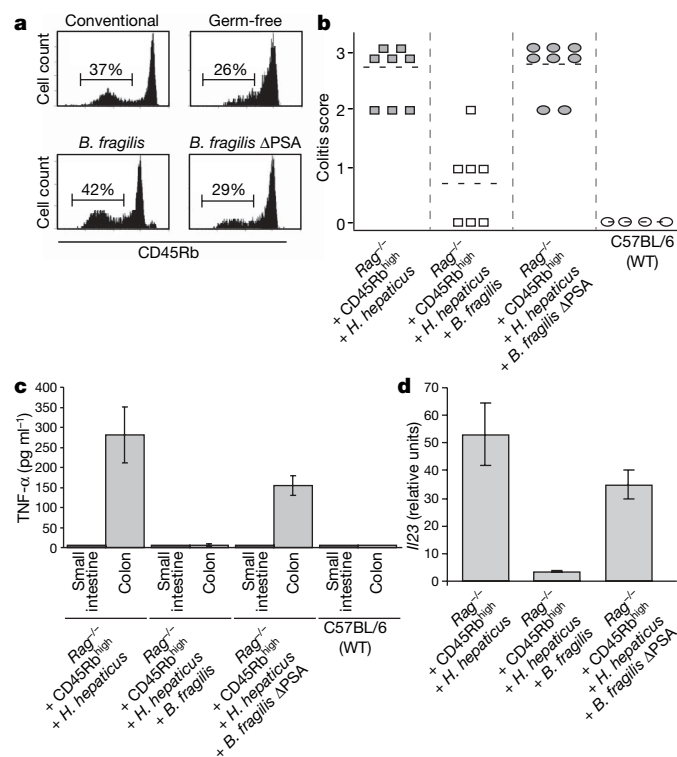


Figure 1 | *B. fragilis* colonization requires PSA for protection from experimental colitis. **a**, Mono-association of germ-free mice with wild-type *B. fragilis* expands the proportions of CD4⁺CD45Rb^{low} T cells in a PSA-dependent manner (mean ± s.d. for 3 experiments: conventional, 38.4% ± 2.2; germ-free, 26.7% ± 1.3; *B. fragilis*, 40.8% ± 3.1; *B. fragilis* ΔPSA, 28.8% ± 2.6). All cells were gated on CD4⁺ splenocytes. **b**, Co-colonization with *B. fragilis* ameliorates disease onset, whereas co-colonization with *B. fragilis* ΔPSA is not protective ($P = 0.004$, Mann–Whitney U -test). Combined data from two independent experiments are shown. **c**, ELISA of colon organ cultures demonstrates increased expression of pro-inflammatory cytokine TNF- α in diseased colons, with significant reductions in animals co-colonized with wild-type *B. fragilis* but not with *B. fragilis* ΔPSA. **d**, qPCR for *Il23p19* was performed on splenocytes and normalized to *L32* expression. Error bars represent s.d. for triplicate samples in all cases. WT, wild type.

The pro-inflammatory cytokines tumour-necrosis factor- α (TNF- α), interleukin (IL)-1 β and IL-23 are central to disease initiation and progression in this experimental model of colitis²³. Furthermore, amounts of these cytokines are elevated in patients with IBD²⁴, and therapies neutralizing TNF- α have yielded promising results in clinical trials in patients with Crohn's disease²⁵. We examined inflammatory cytokine concentrations during disease by directly culturing intestinal tissues of T-cell recipient colonized animals²⁶. Amounts of the pro-inflammatory cytokines TNF- α (Fig. 1c), IL-12p40 and IL-1 β (Supplementary Fig. 1) are elevated in the colons of *Rag*^{−/−} recipient mice colonized with *H. hepaticus*, but not in sections of small intestine—a site that is not affected in this model. Consistent with the protection observed by pathophysiological analysis of experimental colitis, TNF- α production was not elevated when these animals were co-colonized with wild-type *B. fragilis*. T-cell transfer plus co-colonization with *H. hepaticus* and *B. fragilis* ΔPSA results in increased colonic cytokine production similar to that seen in *Rag*^{−/−} animals colonized with *H. hepaticus* alone. Moreover, purified splenic CD4⁺ T cells from *H. hepaticus*-colonized animals demonstrate increased TNF- α production; this condition is corrected by colonization with wild-type *B. fragilis* but not with the PSA deletion strain (Supplementary Fig. 2). Expression of IL-23 is critical in the cascade of events leading to experimental colitis^{27,28}. We found that increases in *Il23* expression by splenocytes after disease induction are completely suppressed by intestinal colonization with PSA-producing *B. fragilis* (Fig. 1d). Over the course of the experiments, amounts of *H. hepaticus* and *B. fragilis* colonization did not differ between groups; thus, protection is not the result of bacterial clearance (Supplementary Figs 3 and 4). Instead, *B. fragilis* has evolved a specific immunomodulatory molecule that orchestrates beneficial immune responses to prevent its host from developing colitis.

Purified PSA prevents gut pathology

To determine whether PSA is sufficient for protection from disease in the absence of the intact organism, we purified PSA to homogeneity²⁹ and administered it by gavage (orally) to *Rag*^{−/−} mice. We measured disease progression by various pathological and histological criteria. Colitis scores after CD4⁺CD45Rb^{high} T-cell transfer in the absence of *H. hepaticus* colonization indicate the development of very mild colitis due to inflammation elicited by the animals' specific pathogen-free microbiota (Fig. 2a; first column). However, *H. hepaticus*-colonized *Rag*^{−/−} animals that received CD4⁺CD45Rb^{high} T-cell transfers developed severe colitis (Fig. 2a; second column). Oral PSA administration almost completely protected animals against *H. hepaticus*-induced colitis (Fig. 2a; third column), reducing disease to levels of control animals which do not develop colitis (Fig. 2a; fourth column).

The inability to gain weight is a hallmark of colitis in this experimental setting⁴. Wasting disease occurred in *Rag*^{−/−} animals after transfer of CD4⁺CD45Rb^{high} cells and colonization with *H. hepaticus* (Fig. 2b; PBS + *Hh*). These animals also developed intestinal pathology and expressed pro-inflammatory cytokines (as described above). Oral administration of PSA from the outset completely protected animals against *H. hepaticus*-mediated wasting disease (Fig. 2b; PSA + *Hh*). Demonstrating that *Helicobacter hepaticus* provides the necessary antigens for inflammation induction, no pathology was observed in uncolonized animals (Fig. 2b; PBS − *Hh*) or in animals without cell transfer.

We examined histological sections of colons for inflammation resulting in experimental colitis. Transfer of CD4⁺CD45Rb^{high} T cells into *H. hepaticus*-colonized *Rag*^{−/−} mice resulted in onset of severe colitis, as evidenced by massive epithelial cell hyperplasia and gross thickening of the gut wall (Fig. 2c; second panel). Consistent with previous studies, the combination of CD4⁺CD45Rb^{high} T-cell transfer and *H. hepaticus* colonization resulted in leukocyte infiltration of affected tissues—a hallmark of inflammation and disease

(Fig. 2c; second panel, bottom)^{19,21}. Oral administration of PSA into *H. hepaticus*-colonized cell transfer recipients conferred complete protection against experimentally induced colonic hyperplasia (Fig. 2c; third panel). Furthermore, PSA-treated animals showed no leukocyte infiltration in colonic tissues (Fig. 2c; third panel, bottom), indicating protection against inflammation. Taken together, these results suggest that PSA prevents colitis and protects mice against the associated weight loss and inflammatory cell infiltration observed in diseased animals.

Control of chemically induced inflammation

Experimental colitis and human IBD result from an initial inflammatory response that—lacking repression—advances in an uncontrolled fashion and ultimately leads to intestinal pathology and disease. To elucidate how PSA affects these primary inflammatory responses, we used an animal model of chemically induced colonic inflammation. Rectal administration of trinitrobenzene sulphonic acid (TNBS) to wild-type mice mimics the initiation of colitis by

eliciting inflammatory T-cell responses. Disease was induced by the administration of TNBS (vehicle was used as a negative control), and oral treatment of PSA was evaluated. TNBS-treated animals exhibited the greatest amount of weight loss, and were unable to recover as rapidly in comparison to either vehicle-treated or PSA-treated animals. (Fig. 3a). Histological analysis confirmed PSA protection of colonic tissues against the massive epithelial hyperplasia and loss of colonocyte organization seen after TNBS treatment (Fig. 3b). Studies have shown that pathogenic T-helper (T_H)17 cells, which produce IL-17, mediate the induction of TNBS experimental colitis³⁰. We found that *Il17* expression was increased among purified $CD4^+$ T cells from mesenteric lymph nodes (MLNs; Fig. 3c) of diseased animals but not from those receiving PSA treatment. The increased expression of *Tnfa* among $CD4^+$ T cells from MLNs of TNBS-treated animals was also reduced in PSA-treated groups (Fig. 3d). Transcriptional analysis of TNBS-treated colons demonstrated that the expression of both *Il17* and *Tnfa* was highly elevated in diseased but not in PSA-protected animals (Fig. 3e, f). Therefore, PSA inhibits intestinal pathology and inflammation in a chemically induced model of experimental colitis.

PSA induces production of IL-10

Protection from experimental colitis is engendered through anti-inflammatory processes that prevent undesirable reactions against the intestinal microbiota²³. Interleukin-10-deficient (*Il10*^{-/-}) animals develop colitis³¹. IL-10 is one of the most potent anti-inflammatory cytokines and is required for protection in many animal models of inflammation^{21,27,32}. As assayed by quantitative real-time polymerase chain reaction (qPCR), transcriptional expression of *Il10* within colons of PSA-treated mice was significantly higher than in control and TNBS-treated mice (Fig. 4a). IL-10 is produced by many cell types; however, $CD4^+$ T cells that express IL-10 have immunosuppressive activities that inhibit inflammation during experimental colitis³³. We purified fresh $CD4^+$ T cells from MLNs of PSA-treated mice in which inflammation was reduced, and we found greatly elevated expression of the *Il10* transcript (Fig. 4b). We assessed whether PSA was sufficient to induce IL-10 *in vitro*; a specific increase in IL-10 production occurred when bone-marrow-derived dendritic cells (BMDCs) and naive $CD4^+$ T cells were treated with purified PSA (Fig. 4c). When BMDCs and naive $CD4^+$ T cells were infected with *H. hepaticus* co-cultured with *B. fragilis*, we found specific expression of IL-10 from culture supernatants, whereas co-culture with *B. fragilis* Δ PSA induced significantly lower expression of IL-10 (Supplementary Fig. 5). As PSA induces expression of IL-10 *in vitro*, we speculated whether this molecule is required for inhibition of inflammatory responses to *H. hepaticus*. We infected BMDC-T-cell co-cultures with live *H. hepaticus* and measured production of

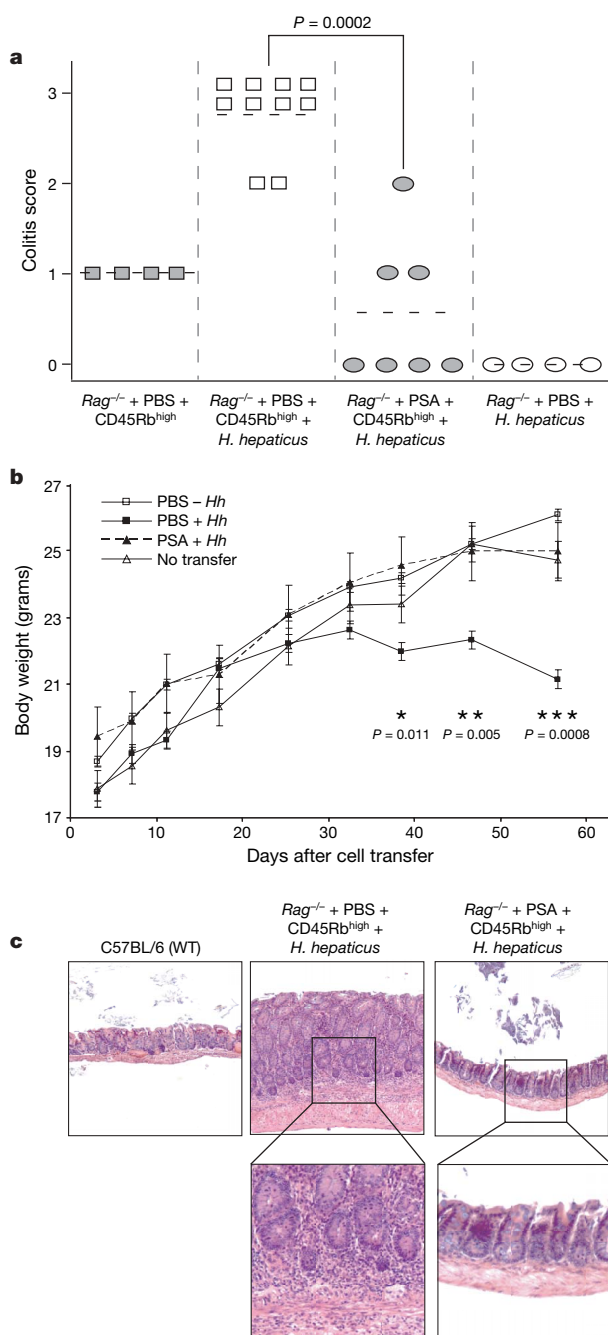


Figure 2 | Purified PSA protects against experimental colitis.

a, Colonization with *H. hepaticus* drives the onset of severe colitis (second column). PSA-treated animals either do not develop colitis or develop only mild disease (third column) ($P > 0.05$ for dissimilar results, $P < 0.01$ for similar results, Kruskal–Wallis comparison of all groups; $P = 0.0002$, two-tailed Mann–Whitney *U*-test). The data shown are the pooled results from duplicate experiments. **b**, Wasting disease in *Rag2*^{-/-} animals results from transfer of $CD4^+$ $CD45Rb^{\text{high}}$ T cells and colonization with *H. hepaticus* (PBS + *Hh*). PSA treatment by gavage protects animals against wasting (PSA + *Hh*). Two-factor analysis of variance (ANOVA) indicates that comparisons between all indicated groups (asterisks) are statistically significant. Error bars represent s.d. between four animals per group; experiments were performed in duplicate. **c**, Architecture of colonic sections from wild-type animals (first panel). $CD4^+$ $CD45Rb^{\text{high}}$ T-cell transfer into *H. hepaticus*-colonized *Rag2*^{-/-} mice resulted in severe colitis, as evidenced by massive epithelial hyperplasia and pronounced inflammation (second panel); the higher magnification below shows inflammatory cell infiltration into colonic tissues. Oral PSA treatment protects *H. hepaticus*-colonized animals (third panel) from colitis. Images in each row are the same magnification; original magnification $\times 10$ for top, $\times 40$ for bottom row. WT, wild type.

the pro-inflammatory cytokine TNF- α . Addition of increasing amounts of the pathogenic commensal bacterium resulted in a dose-dependent increase in TNF- α concentration, as determined by enzyme-linked immunosorbent assay (ELISA) of culture supernatants (Fig. 4d; left three bars). Treatment of cells with purified PSA decreased TNF- α production in response to *H. hepaticus* (Fig. 4d; middle three bars). Most notably, co-incubation of cell cultures with *H. hepaticus* and PSA in the presence of a neutralizing IL-10 receptor antibody completely reversed this phenotypic effect and increased expression of TNF- α (Fig. 4d; right three bars). The results are similar for the related pro-inflammatory cytokine IL-1 β (Supplementary Fig. 6). Thus, IL-10 produced in response to PSA is required for inhibition of inflammatory reactions in cell cultures.

IL-10-producing T cells suppress colitis

We investigated the requirement for IL-10 in suppression of intestinal inflammation. Initially, *Il10*^{-/-} animals were colonized with *H. hepaticus* alone or in combination with *B. fragilis* (either wild type or Δ PSA). We subsequently collected MLNs and re-stimulated cells in culture with soluble *H. hepaticus* antigens using an assay previously developed to measure antigen-specific responses to this pathogen²⁷. *H. hepaticus*-colonization resulted in increased production of TNF- α

and IL-17 by MLN cells; however, in the absence of IL-10 production in colonized animals, *B. fragilis* co-colonization did not reduce the concentration of these pro-inflammatory molecules (Fig. 5a and b, respectively). As expected, the absence of PSA had no effect. Again using the cell transfer model of colitis, we transferred CD4⁺CD45Rb^{high} T cells to *H. hepaticus*-colonized *Rag*^{-/-} animals. Administration of an anti-IL-10 receptor antibody to mice (to block IL-10 signalling) during oral treatment with PSA abrogated protection from colitis (Fig. 5c). When *Il10*^{-/-} animals were treated with TNBS in the presence or absence of PSA, weight and histology data (Supplementary Figs 7 and 8) indicate that IL-10 production is required for PSA-elicited reduction of intestinal immune responses.

Our data suggest that PSA-mediated protection entails the generation and/or expansion of IL-10-producing CD4⁺ T cells. To determine whether IL-10 production by CD4⁺ T cells is required for protection, we transferred CD4⁺CD45Rb^{high} T cells from *Il10*^{-/-} donor mice into *Rag*^{-/-} recipients and then colonized the recipients with *H. hepaticus*. As expected, groups of mice receiving *Il10*^{-/-} T cells along with *H. hepaticus* developed severe colitis (Fig. 5d; first column) and were not protected by PSA (Fig. 5d; second column). This result, supported by histological findings in colons, suggests that PSA induces protection from 'previously pathogenic' CD4⁺CD45Rb^{high} T cells in an IL-10-dependent manner (Fig. 5e). Weight analysis at death shows that colitic PBS- and PSA-treated animals receiving *Il10*^{-/-} CD4⁺CD45Rb^{high} T cells (unlike control animals receiving no transferred cells) developed wasting disease (Fig. 5f). Thus, IL-10 production by CD4⁺ T cells is required for PSA-mediated protection from experimental colitis. These results constitute the first reported evidence of a symbiotic bacterial molecule that networks with the immune system to coordinate anti-inflammatory responses required for mammalian health.

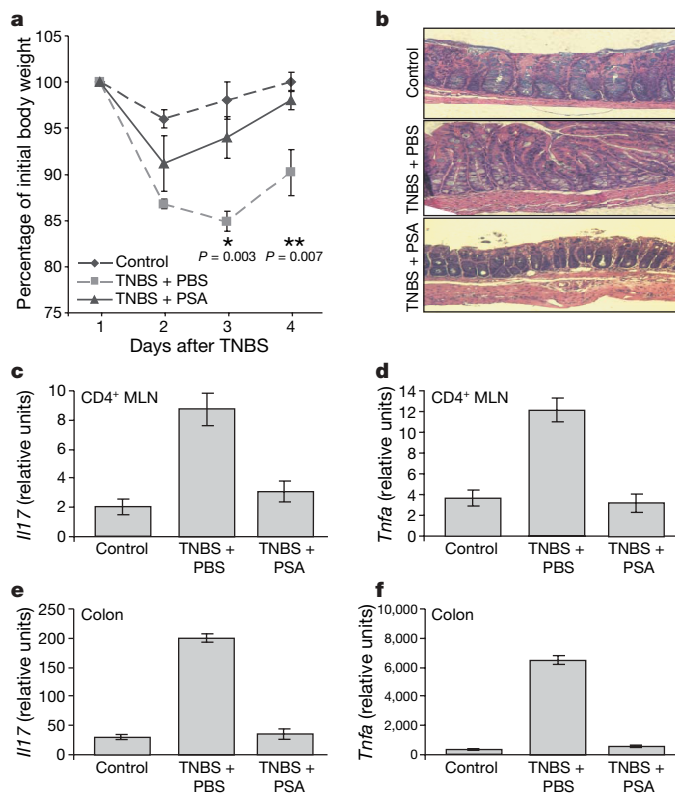


Figure 3 | Intestinal immune responses are modulated by a beneficial bacterial molecule. **a**, Oral PSA administration causes a statistically significant increase in body weight related to TNBS-treated PBS controls. Two-factor ANOVA values for all indicated groups (asterisks) are statistically significant. Error bars represent s.d. between four animals per group. **b**, Colons from TNBS plus PBS-treated groups show severe pathology, whereas those from TNBS and PSA-treated animals have histological architecture similar to that seen in untreated controls. The images shown are representative sections from animals in two independent experiments. Original magnification $\times 20$. **c**, **d**, qPCR of purified CD4⁺ T cells from MLNs with *Il17a*- and *Tnfa*-specific primers demonstrates that oral PSA administration reduces *Il17a* (**c**) and *Tnfa* (**d**) expression during disease. **e**, **f**, Transcriptional expression of *Il17a* (**e**) and *Tnfa* (**f**) from homogenized colons. Error bars are from duplicate runs of three independent experiments (**c**–**f**).

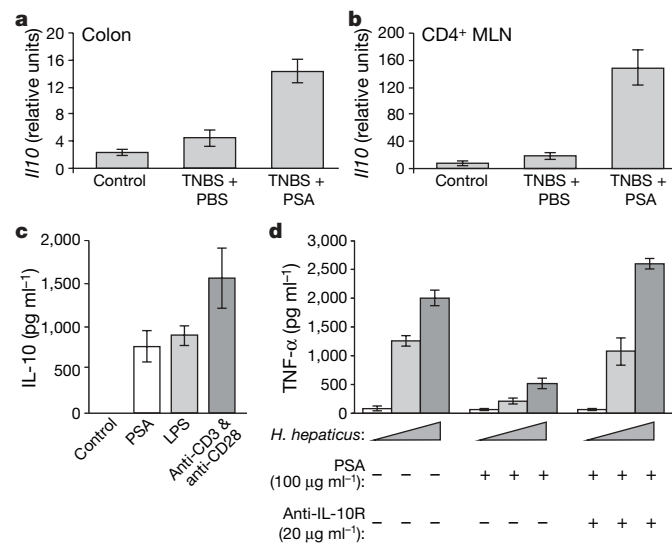


Figure 4 | PSA induces *Il10* expression in TNBS-treated animals and inhibits TNF- α production in primary cultured cells through IL-10 production. **a**, **b**, Wild-type mice were treated with ethanol (control), TNBS, or TNBS and PSA. qPCR assay of colons (**a**) and CD4⁺ T cells purified from MLNs (**b**) show elevated *Il10* expression in response to PSA. **c**, Incubation of BMDC-T cell co-cultures with purified PSA specifically induces IL-10 production at concentrations comparable to those induced by LPS or by anti-CD3 and anti-CD28 antibodies. **d**, Infection of BMDC-T cell co-cultures with increasing concentrations of *H. hepaticus* (multiplicity of infection: 0.1, 1.0 and 10, as depicted by triangles) results in increased TNF- α release. Treatment of infected cells with PSA reduces amount of TNF- α (middle three bars). Addition of an anti-IL-10 receptor antibody (anti-IL-10R) alleviates suppression of inflammatory responses, resulting in increased TNF- α production (right three bars). Error bars show s.d. for triplicate samples in all cases.

Summary and implications

According to the 'hygiene hypothesis' put forth nearly two decades ago, reduced exposure to infections in early childhood—owing to diminishing family size and improvements in living standards and personal hygiene, for example—may increase the risk of allergic and autoimmune disease³⁴. This concept is supported by epidemiological and clinical reports documenting increased incidences of IBD, colon cancer, asthma, type 1 diabetes and multiple sclerosis over the past 50 years in societies with improved medical care and hygiene (for example, Europe, the United States and Japan) but not in undeveloped countries³⁵. However, the application of major interventions, including vaccination, sanitation, and antibacterial and antiviral therapies, often does not permit discrimination between

infectious and non-infectious microorganisms and has undoubtedly led to changes in human association with the microbial world as a whole. The hygiene hypothesis does not address humanity's primary relationship with bacteria: the harbouring of multitudes of microbial species during commensalism. Our studies show that symbiotic bacteria residing in the mammalian gastrointestinal tract produce molecules that mediate healthy immune responses and protect the host from inflammatory disease. We propose that the mammalian genome does not encode for all functions required for immunological development but rather that mammals depend on critical interactions with their microbiome (the collective genomes of the microbiota) for health.

As mammals have harboured their commensal partners for millennia, adaptive co-evolution has formed an inextricable bond between microbe and host³⁶. Imbalances in the microbiota may contribute to some human diseases, and altered composition of the gut bacteria has been implicated in obesity³⁷. We show that *B. fragilis* protects its host from inflammatory disease caused by *H. hepaticus* in an animal model of experimental colitis. The implication that intestinal bacteria actively network with the host's immune system highlights the importance of the composition of the microbiota for overall health. If specific classes of bacteria have indeed evolved to promote the host's health, then disease may well result from the absence of these organisms and their beneficial molecules (for example, as a result of improved hygiene). Inflammation resulting from dysbiosis between symbionts and pathobionts may lay the molecular foundations for many intestinal—and perhaps non-intestinal—diseases. The exploration of probiotics (bacteria such as lactobacilli and bifidobacteria that promote health) has thus far failed to identify specific bacterial molecules or host mechanisms required for protection³⁸. Here we present evidence that a single bacterial molecule can ameliorate inflammatory disease in animals. Our observations suggest that many other symbiosis factors—bacterial molecules that have evolved to promote human health—remain to be discovered. The finding that PSA from *B. fragilis* is a natural anti-inflammatory molecule that actively promotes mammalian health may provide a platform for the development of therapies based on the fundamental relationship between humans and their beneficial microbial partners.

METHODS SUMMARY

Three models of intestinal inflammation were used: (1) CD4⁺CD45Rb^{high} T cells were purified from the spleens of wild-type or *Il10*^{-/-} donor mice by flow cytometry and transferred into *Rag*^{-/-} (C57BL/6) recipients as described; (2) TNBS colitis was induced by pre-sensitization of wild-type (C57BL/6) mice on the skin with a TNBS and acetone mix. Seven days after sensitization, 2.5% TNBS in ethanol was administered rectally; mice were killed 3–6 days later; and (3) *Il10*^{-/-} mice were colonized (by oral gavage) with *H. hepaticus* alone or in combination with wild-type *B. fragilis* or *B. fragilis* ΔPSA. *B. fragilis* NCTC 9343 and *H. hepaticus* ATCC 51149 were obtained from the American Type Culture Collection. Cytokines from the spleens, colons, or MLNs were assayed by ELISA, qPCR, or flow cytometry. Colitis was assessed with tissue sections (fixed, paraffin embedded, sectioned onto a slide and stained with haematoxylin and eosin) and was scored by a pathologist in a blinded experimental set-up (R. T. Bronson) according to a standard scoring system: 0, no thickening of colonic tissues and no inflammation (infiltration of lymphocytes); 1, mild thickening of tissues but no inflammation; 2, mild thickening of tissues and mild inflammation; 3, severe thickening and severe inflammation. BMDCs were purified from femurs of mice after extraction and washing in PBS. Cells were cultured for 8 days in C-RPMI-10 in the presence of GM-CSF (20 ng ml⁻¹; Biosource). CD4⁺ T cells were purified by negative selection over a magnetic column (Miltenyi or R&D Systems).

Received 7 February; accepted 18 April 2008.

- Poxton, I. R., Brown, R., Sawyerr, A. & Ferguson, A. Mucosa-associated bacterial flora of the human colon. *J. Med. Microbiol.* **46**, 85–91 (1997).
- Sellon, R. K. et al. Resident enteric bacteria are necessary for development of spontaneous colitis and immune system activation in interleukin-10-deficient mice. *Infect. Immun.* **66**, 5224–5231 (1998).

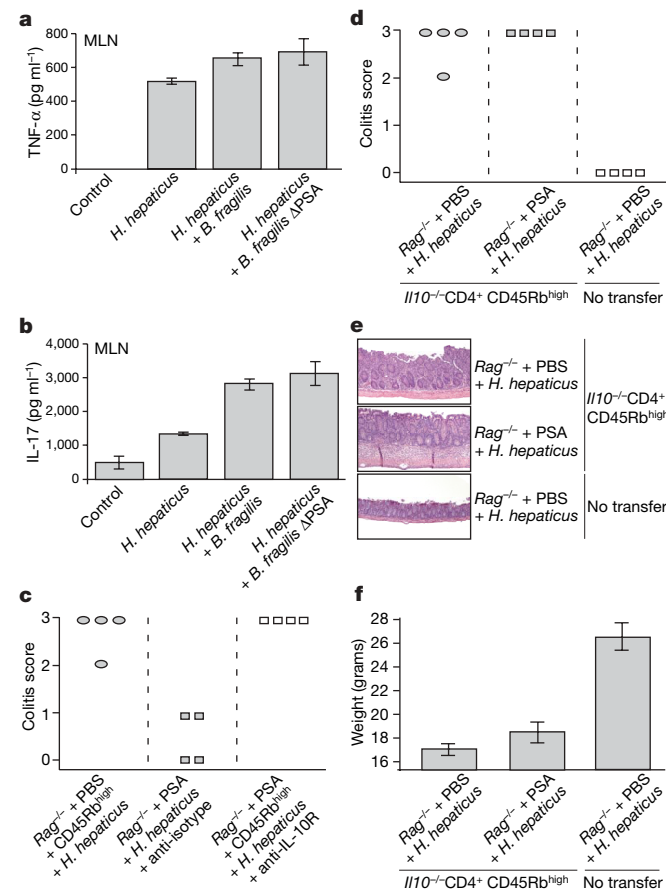


Figure 5 | IL-10 is required for PSA-mediated protection from intestinal inflammation and experimental colitis. **a, b**, *Il10*^{-/-} mice were left uncolonized (control) or were colonized with *H. hepaticus* (to induce inflammation) either alone or in combination with *B. fragilis* (wild-type or ΔPSA). MLNs from experimental groups were pooled and re-stimulated with soluble *H. hepaticus* antigen (5 μg ml⁻¹) for 48 h. Secretion of the pro-inflammatory cytokines TNF-α (**a**) and IL-17A (**b**) was analysed by ELISA. Error bars show s.d. for triplicate samples. **c**, Colitis scores show that PSA protection requires IL-10 signalling, as treatment with anti-IL-10 receptor antibody (anti-IL-10R) abrogates the PSA-mediated protection. Data represent two independent experiments. **d**, PSA-mediated protection from disease requires IL-10-producing CD4⁺ T cells. Treatment with PSA does not reduce colitis when CD4⁺CD45Rb^{high} T cells are transferred from *Il10*^{-/-} mice. Control animals without cell transfer do not develop colitis. Results are shown for one representative trial of two independent experiments. **e**, Histological colon sections show that PSA does not protect animals from experimental colitis when CD4⁺CD45Rb^{high} T cells cannot produce IL-10. All images are from the same magnification. Original magnification ×10. **f**, Mean body weights for groups of animals (*n* = 4) when culled demonstrate that IL-10 is required for PSA-mediated protection from wasting. Error bars show s.d. between 4 animals per group from one representative trial of two independent experiments.

3. Elson, C. O. Commensal bacteria as targets in Crohn's disease. *Gastroenterology* **119**, 254–257 (2000).
4. Sartor, R. B. Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis. *Nature Clin. Pract. Gastroenterol. Hepatol.* **3**, 390–407 (2006).
5. Videla, S. *et al.* Role of intestinal microflora in chronic inflammation and ulceration of the rat colon. *Gut* **35**, 1090–1097 (1994).
6. Taurog, J. D. *et al.* The germfree state prevents development of gut and joint inflammatory disease in HLA-B27 transgenic rats. *J. Exp. Med.* **180**, 2359–2364 (1994).
7. Kullberg, M. C. *et al.* Induction of colitis by a CD4⁺ T cell clone specific for a bacterial epitope. *Proc. Natl Acad. Sci. USA* **100**, 15830–15835 (2003).
8. O'Hara, A. M. & Shanahan, F. The gut flora as a forgotten organ. *EMBO Rep.* **7**, 688–693 (2006).
9. Frank, D. N. *et al.* Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc. Natl Acad. Sci. USA* **104**, 13780–13785 (2007).
10. Gill, S. R. *et al.* Metagenomic analysis of the human distal gut microbiome. *Science* **312**, 1355–1359 (2006).
11. Ley, R. E., Peterson, D. A. & Gordon, J. I. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* **124**, 837–848 (2006).
12. Smith, K., McCoy, K. D. & Macpherson, A. J. Use of axenic animals in studying the adaptation of mammals to their commensal intestinal microbiota. *Semin. Immunol.* **19**, 59–69 (2006).
13. Mazmanian, S. K., Liu, C. H., Tzianabos, A. O. & Kasper, D. L. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* **122**, 107–118 (2005).
14. Pamer, E. G. Immune responses to commensal and environmental microbes. *Nature Immunol.* **8**, 1173–1178 (2007).
15. Dethlefsen, L., McFall-Ngai, M. & Relman, D. A. An ecological and evolutionary perspective on human–microbe mutualism and disease. *Nature* **449**, 811–818 (2007).
16. Bell, E. B. Function of CD4 T cell subsets *in vivo*: expression of CD45R isoforms. *Semin. Immunol.* **4**, 43–50 (1992).
17. Izcue, A., Coombes, J. L. & Powrie, F. Regulatory T cells suppress systemic and mucosal immune activation to control intestinal inflammation. *Immunol. Rev.* **212**, 256–271 (2006).
18. Maloy, K. J. *et al.* CD4⁺CD25⁺ T_R cells suppress innate immune pathology through cytokine-dependent mechanisms. *J. Exp. Med.* **197**, 111–119 (2003).
19. Cahill, R. J. *et al.* Inflammatory bowel disease: an immunity-mediated condition triggered by bacterial infection with *Helicobacter hepaticus*. *Infect. Immun.* **65**, 3126–3131 (1997).
20. Scheinin, T., Butler, D. M., Salway, F., Scallan, B. & Feldmann, M. Validation of the interleukin-10 knockout mouse model of colitis: antitumour necrosis factor-antibodies suppress the progression of colitis. *Clin. Exp. Immunol.* **133**, 38–43 (2003).
21. Kullberg, M. C. *et al.* Bacteria-triggered CD4⁺ T regulatory cells suppress *Helicobacter hepaticus*-induced colitis. *J. Exp. Med.* **196**, 505–515 (2002).
22. Bregenholt, S. Cells and cytokines in the pathogenesis of inflammatory bowel disease: new insights from mouse T cell transfer models. *Exp. Clin. Immunogenet.* **17**, 115–129 (2000).
23. Powrie, F. & Maloy, K. J. Immunology. Regulating the regulators. *Science* **299**, 1030–1031 (2003).
24. Xavier, R. & Podolsky, D. K. Commensal flora: wolf in sheep's clothing. *Gastroenterology* **128**, 1122–1126 (2005).
25. Rutgeerts, P. *et al.* Infliximab for induction and maintenance therapy for ulcerative colitis. *N. Engl. J. Med.* **353**, 2462–2476 (2005).
26. Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S. & Medzhitov, R. Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell* **118**, 229–241 (2004).
27. Kullberg, M. C. *et al.* IL-23 plays a key role in *Helicobacter hepaticus*-induced T cell-dependent colitis. *J. Exp. Med.* **203**, 2485–2494 (2006).
28. Hue, S. *et al.* Interleukin-23 drives innate and T cell-mediated intestinal inflammation. *J. Exp. Med.* **203**, 2473–2483 (2006).
29. Tzianabos, A. O. *et al.* The capsular polysaccharide of *Bacteroides fragilis* comprises two ionically linked polysaccharides. *J. Biol. Chem.* **267**, 18230–18235 (1992).
30. Elson, C. O. *et al.* Monoclonal anti-interleukin 23 reverses active colitis in a T cell-mediated model in mice. *Gastroenterology* **132**, 2359–2370 (2007).
31. Kuhn, R., Lohler, J., Rennick, D., Rajewsky, K. & Muller, W. Interleukin-10-deficient mice develop chronic enterocolitis. *Cell* **75**, 263–274 (1993).
32. Asseman, C., Mauze, S., Leach, M. W., Coffman, R. L. & Powrie, F. An essential role for interleukin 10 in the function of regulatory T cells that inhibit intestinal inflammation. *J. Exp. Med.* **190**, 995–1004 (1999).
33. Groux, H. *et al.* A CD4⁺ T-cell subset inhibits antigen-specific T-cell responses and prevents colitis. *Nature* **389**, 737–742 (1997).
34. Strachan, D. P. Hay fever, hygiene, and household size. *Br. Med. J.* **299**, 1259–1260 (1989).
35. Bach, J. F. The effect of infections on susceptibility to autoimmune and allergic diseases. *N. Engl. J. Med.* **347**, 911–920 (2002).
36. Liu, C. H., Lee, S. M., Vanlare, J. M., Kasper, D. L. & Mazmanian, S. K. Regulation of surface architecture by symbiotic bacteria mediates host colonization. *Proc. Natl Acad. Sci. USA* **105**, 3951–3956 (2008).
37. Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027–1031 (2006).
38. Mazmanian, S. K. & Kasper, D. L. The love-hate relationship between bacterial polysaccharides and the host immune system. *Nature Rev. Immunol.* **6**, 849–858 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. T. Bronson for discussions about histopathology; members of the Mazmanian laboratory for critical comments throughout the course of the work; and J. McCoy for editorial expertise. S.K.M. acknowledges a fellowship from the Helen Hay Whitney Foundation; J.L.R. acknowledges support from the Jane Coffin Childs Memorial Fund. This work was supported by funding from the NIH/NIAID (R01 AI039576) to D.L.K., and by grants from the Searle Scholars Program, the Damon Runyon Cancer Research Foundation, and the Crohn's and Colitis Foundation of America to S.K.M.

Author Contributions S.K.M., J.L.R. and D.L.K. designed the research; S.K.M. and J.L.R. performed the research; S.K.M., J.L.R. and D.L.K. analysed the data and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.K.M. (sarkis@caltech.edu) or D.L.K. (dennis_kasper@hms.harvard.edu).

LETTERS

An infrared ring around the magnetar SGR 1900+14

S. Wachter¹, E. Ramirez-Ruiz², V. V. Dwarkadas³, C. Kouveliotou⁴, J. Granot⁵, S. K. Patel⁶ & D. Figer⁷

Magnetars^{1,2} are a special class of slowly rotating (period $\sim 5\text{--}12$ s) neutron stars with extremely strong magnetic fields ($>10^{14}$ G)—at least an order of magnitude larger than those of the ‘normal’ radio pulsars. The potential evolutionary links and differences between these two types of object are still unknown; recent studies, however, have provided circumstantial evidence connecting magnetars with very massive progenitor stars^{3–5}. Here we report the discovery of an infrared elliptical ring or shell surrounding the magnetar SGR 1900+14. The appearance and energetics of the ring are difficult to interpret within the framework of the progenitor’s stellar mass loss or the subsequent evolution of the supernova remnant. We suggest instead that a dust-free cavity was produced in the magnetar environment by the giant flare emitted by the source in August 1998. Considering the total energy released in the flare, the theoretical dust-destruction radius matches well with the observed dimensions of the ring. We conclude that SGR 1900+14 is unambiguously associated with a cluster of massive stars, thereby solidifying the link between magnetars and massive stars.

Soft gamma repeaters (SGRs) and anomalous X-ray pulsars (AXPs) are the two main classes of objects currently believed to be magnetars. Their characteristic ages, derived from their rotational properties⁶, indicate a young population—typically a few thousand years old. Although these spin-down ages must be treated with caution, the relative youth of SGRs and AXPs is supported by the fact that some AXPs are located at the centres of supernova remnants (SNRs). Associations between SNRs and SGRs have also been claimed in the literature, but, unlike the AXPs, the SGRs are offset from the centres of their proposed SNRs, increasing the likelihood of spurious alignment. As a result, the validity of all SGR–SNR associations has been questioned⁷. Both AXPs and SGRs, for example the AXP CXOU J164710.2–455216 (ref. 5) and SGR 1806–20 (ref. 8), have been linked to clusters of massive stars. The implied progenitor mass of $M \geq 40\text{--}50 M_{\odot}$ supports theoretical predictions that very massive stars with sufficient metallicity and the corresponding high mass loss can still form neutron stars⁹, contrary to the standard evolutionary picture that such massive stars end their lives as black holes.

As part of our systematic study of the circumstellar environments of SGRs and AXPs, we observed the position of SGR 1900+14 using all three instruments onboard the NASA Spitzer Space Telescope in 2005 and 2007. Surprisingly, our 24- μm and 16- μm images (Fig. 1b, c) reveal a prominent ring-like structure that is not detected in our 3.6–8.0- μm observations. A formal elliptical fit to the ring indicates semi-major and semi-minor axes of angular lengths $\sim 36''$ and $\sim 19''$, respectively, centred at right ascension 19 h 07 min 14.32 s and declination $+09^{\circ} 19' 20.0'' \pm 1.0''$. This coincides with the radio position of SGR 1900+14 (19 h 07 min 14.33 s, $+09^{\circ} 19' 20.1'' \pm 0.15''$), determined from observations of a transient synchrotron source associated with the 1998 giant flare¹⁰. We re-examined

published and archival data of the field around SGR 1900+14 but found no equivalent structure at optical, near-infrared, radio or X-ray wavelengths. In particular, the lack of detection in the radio spectrum, to a limit of $L_{332\text{ MHz}} \leq 4 \times 10^{29} d_{15}^2 \text{ erg s}^{-1}$, and the X-ray spectrum, to a limit of $L_{2\text{--}10\text{ keV}} \leq 2.7 \times 10^{33} d_{15}^2 \text{ erg s}^{-1}$ (Chandra Data Archive, <http://cxc.harvard.edu/cda/>), is important in determining the nature of the ring. Here d_{15} parameterizes the distance to the source, according to $d = 15 d_{15} \text{ kpc}$.

The Spitzer images are dominated by the bright emission from two nearby M5 supergiants that mark the centre of a compact cluster of massive stars at a distance of 12–15 kpc^{12,13}. Although it has been suggested that SGR 1900+14 is associated with this cluster¹², an alternative distance of 5 kpc has also been suggested, on the basis of the hydrogen column density of its X-ray spectrum¹⁴. In addition, the SGR is offset by $12''$ from the cluster centre, and the visual extinction corresponding to the X-ray-derived hydrogen column density¹¹ ($A_V = 12.8 \pm 0.8 \text{ mag}$) is significantly different from that deduced optically for the cluster stars¹² ($A_V = 19.2 \pm 1 \text{ mag}$). Considering both possibilities, we calculated physical sizes for the semi-major and semi-minor axes of the ring of $0.9 \times 0.5 \text{ pc}$, if SGR 1900+14 lies at a distance of $\sim 5 \text{ kpc}$, or $2.6 \times 1.4 \text{ pc}$, if it resides at a cluster distance of 15 kpc.

We have constrained the temperature of the material in the ring by first cross-convolving the images with the point spread function of the different arrays. We then measured the flux through fixed apertures at several positions along the ring. The largest source of uncertainty lies in the subtraction of the local background. We derived a temperature of 130–150 K for the material in the ring using a simple black-body fit, and one of $\sim 80\text{--}120 \text{ K}$ using a more realistic dust model¹⁵. We caution that the presence of spectral lines or peculiarities in the underlying mid-infrared spectral shape might significantly alter these derived temperatures. For example, Spitzer observations have revealed a shell only visible at 24 μm , where the broadband flux can be entirely attributed to [O IV] line emission¹⁶. A broad 22- μm continuum feature has also been identified in environments associated with supernovae and massive star formation¹⁷.

The 70- μm image shows a bipolar flux distribution along the minor axis of the ellipse (Fig. 1d). Extended enhancements in the ring are also seen at these positions in the 16- μm and 24- μm images. Our data do not allow us to determine the underlying three-dimensional geometry of the structure, that is, to distinguish between a limb-brightened shell, a true ring morphology or a bipolar cavity with equatorial torus. Similarly, an accurate measurement of the total flux in the ring is very difficult because the enveloping diffuse emission (Fig. 1a) coupled with the contribution from the M5 supergiants prohibits a clean separation of the ring emission. Using a narrow elliptical aperture, we measured fluxes of $1.2 \pm 0.2 \text{ Jy}$ and $0.4 \pm 0.1 \text{ Jy}$ at 24 μm and 16 μm , respectively, implying ring luminosities of $\nu L_{\nu}(16 \mu\text{m}) \approx 2 \times 10^{36} d_{15}^2 \text{ erg s}^{-1}$ and $\nu L_{\nu}(24 \mu\text{m}) \approx$

¹Spitzer Science Center, California Institute of Technology, Pasadena, California 91125, USA. ²Astronomy & Astrophysics, 201 Interdisciplinary Sciences Building, Santa Cruz, California 95064, USA. ³Department of Astronomy and Astrophysics, University of Chicago, 5640 South Ellis Avenue, AAC 010c, Chicago, Illinois 60637, USA. ⁴NASA/Marshall Space Flight Center, VP62, NSSTC, 320 Sparkman Drive, Huntsville, Alabama 35805, USA. ⁵Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK. ⁶Optical Sciences Corporation, 6767 Old Madison Pike, Suite 650, Huntsville, Alabama 35806, USA. ⁷Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, 54 Lomb Memorial Drive, Rochester, New York 14623, USA.

$4 \times 10^{36} d_{15}^2 \text{ erg s}^{-1}$ (where ν denotes the frequency at which the luminosity is determined).

To test for ring evolution, we used the publicly available MIPS GAL Spitzer Legacy programme observations of the field that were obtained at an epoch (2005 October 03) earlier than was our own data set. Creating a difference image from the 24- μm data of the two different epochs, we find no discernable change in the size (positional shifts $< 0.5''$) or flux (root mean square $< 1\text{--}2\%$) of the ring emission in the 1.6 yr between the two observations. The stationary nature of the ring strongly disfavours one obvious explanation for the creation of the ring—a light echo, due to reprocessing by the surrounding dust, from the 1998 SGR 1990+14 giant flare. This model predicts a typical fractional change in the ring size and brightness between the two epoch observations of $\sim 23\%$, which is not seen in the data (see Supplementary Information for more details).

Morphologically, the ring resembles the wind-blown bubbles and shells observed around evolved massive stars^{18–20}, such as B supergiants, luminous blue variables and Wolf–Rayet stars. The supersonic wind from the star drives a shock into the surrounding medium, sweeping up the ambient material into a thin, dense, cool shell. The emission from the swept-up material is powered by the luminous central star. Such shells have typical radii of $\sim 2\text{--}10$ pc for Wolf–Rayet stars²⁰ and $0.1\text{--}2.3$ pc for luminous blue variables/supergiants^{18,19}, overlapping the size of our ring. To guard against a simple chance superposition with an unrelated source, we investigated the photometric properties of stars near the centre of the ring. The sources in the immediate vicinity of the SGR position¹¹ are too faint in comparison with the near-infrared luminosities expected for the massive stars discussed above^{21,22}, and thus cannot be physically associated with the ring. For stars with larger offsets from the centre ($\geq 7''$) we cannot rule out a massive star classification based solely on available archival data. However, a situation in which an off-centre star creates an elliptical ring with an unrelated magnetar at its exact centre appears rather contrived.

In the absence of an obvious chance superposition, the close positional coincidence between the SGR and the centre of the ring then strongly indicates that the SGR and the ring are physically connected. One interpretation is that the ring comprises material ejected during the late stages of the SGR progenitor's evolution. However, we know that a supernova explosion occurred, producing the magnetar, and the resulting shock wave is expected to disrupt such a close shell as it interacts with and dissipates into the surrounding material. Some cases are known in which the SNR has over-run a wind-swept shell, but the wind bubble invariably emerges irregular and fragmented in shape after its encounter with the supernova²³. The symmetric and well-defined appearance of our ring thus would imply that the supernova shock wave is still inside the ring. Our simulations show that this would require the supernova to be < 200 yr old, with bright radio and X-ray emission, contrary to observations. The lack of high-energy X-ray and non-thermal radio emission also excludes the possibility that the observed ring is produced by the emission of the blast wave from the progenitor's supernova explosion.

The rejection of these various possibilities raises the issue of the energy source that is powering the ring. We derived a limit for the required luminosity of the heating source by means of radiation balance, using the measured temperature T_{IR} of the dust. If the luminosity L_* is dominated by a source of temperature T_* , then (see Supplementary Information for details)

$$L_* \approx 2.7 \times 10^{40} T_{\text{IR}}^5 T_*^{-1} R^2 \text{ erg s}^{-1}$$

where T_{IR} is measured in units of 100 K, T_* is measured in units of 10^4 K and R , the distance from the central heating source, is measured in units of parsecs.

We note that the persistent X-ray luminosity of the SGR is only $L_{2\text{--}10\text{ keV}} \approx 2.0 \times 10^{35} d_{15}^2 \text{--} 3.5 \times 10^{35} d_{15}^2 \text{ erg s}^{-1} \ll L_*$ and, thus, cannot generate the observed infrared emission from the ring. The only viable heat source consistent with the observed properties of the ring appears to be irradiation from the stars in the nearby cluster. Figure 1a clearly

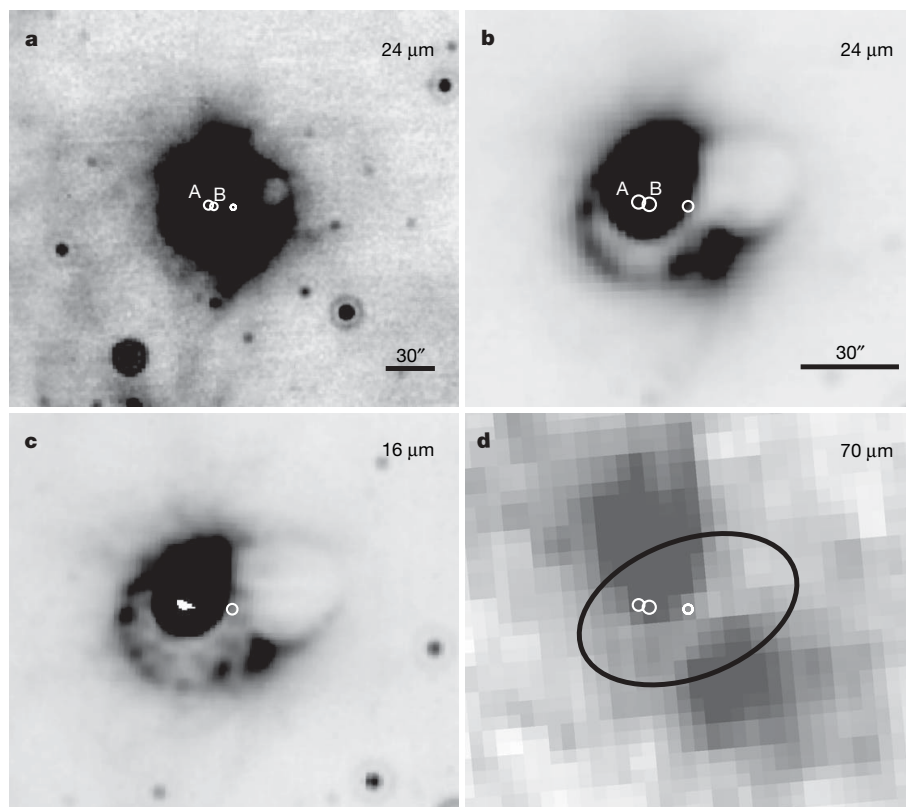


Figure 1 | The infrared ring around SGR 1900+14.

Spitzer Space Telescope imaging of SGR 1900+14 was acquired on 2005 September 20 using the Infrared Array Camera (3.6–8.0 μm), on 2005 October 11 using the blue (16- μm) peak-up imaging mode of the Infrared Spectrograph (IRS), and on 2007 May 29 using the Multiband Imaging Photometer for Spitzer (MIPS; 24 μm , 70 μm). Only the MIPS and IRS images are shown here. We assembled the basic calibrated data products into combined images using the Mosaicking and Point-source Extraction package. For the MIPS data, we first applied a differential flat field to correct for small (1–2%) instrumental artefacts. The final combined images (a–d) have effective exposure times of 31 s for IRS images and 16 s for images from both MIPS channels. The white circle in the centre of each image marks the radio position of the SGR. For all images, north is up and east is to the left. **a**, MIPS 24- μm image scaled to show the diffuse emission enveloping both the SGR and the nearby cluster. The white circles labelled ‘A’ and ‘B’ indicate the positions of the red supergiants¹² located at the centre of the star cluster. **b**, Close-up of the MIPS 24- μm image shown in **a**, but scaled to highlight the ring-like structure of the extended emission. **c**, IRS 16- μm peak-up image covering the same field of view as **b**. **d**, MIPS 70- μm image showing emission associated with the enhanced areas along the minor axis of the ellipse also seen at 16 μm and 24 μm . The position of the ring is indicated by the black ellipsoid. The convolution of the 16- μm and 24- μm images to the resolution of the 70- μm array indicates that the ring emission is too spatially confined and faint to be detectable at this wavelength.

shows that both the cluster and SGR 1900+14 are embedded in a diffuse, extended cloud of 24- μ m emission. Similar diffuse 24- μ m emission is observed in association with the clusters containing SGR 1806–20 (ref. 24) and Westerlund 1, implying that the extended emission is powered by the cluster stars. We note that, given the location and size of the 24- μ m nebula, the cluster might be more spatially extended than previously thought, with as-yet-undiscovered member stars.

A possible mechanism for the creation of the ring is the formation of a dust-free cavity owing to the giant flare activity from SGR 1900+14, which would naturally explain why the ring is centred on the magnetar. The destruction of dust grains can occur either by sublimation due to the heating by ultraviolet²⁵ or X-ray²⁶ emission, or by grain charging due to the incident X-rays, which causes the dust grains to gradually shatter into smaller pieces until they are eventually destroyed²⁶. The size and shape of the ring can constrain the energy output and degree of anisotropy of the flaring event (see Supplementary Information for details). Using equation (25) of ref. 26 with the fiducial values and $\alpha = 0$ (where the flux at frequency ν is proportional to $\nu^{-\alpha}$), we find that the dust-destruction radius matches well with the observed dimensions of the ring around SGR 1900+14 for $E \geq 6 \times 10^{45} d_{15}^2$ erg. This estimate is close to the observed isotropic equivalent energy²⁷ in the initial spike of the 27 August 1998 flare; the latter, however, is only accurate for photon energies ≥ 50 keV, whereas the estimated dust-destruction radius requires photons with energies of ~ 1 keV.

Alternatively, a previous, more energetic giant flare from SGR 1900+14, for example, similar to the 2004 December 27 giant flare from SGR 1806–20 (refs 27, 28), with $E \approx 10^{47} d_{15}^2$ erg, could have carved the dust-free cavity during the magnetar's estimated spin-down age of $t_{\text{age}} \approx 1,800$ yr. For a rate of one flare per ~ 50 yr (ref. 28), it is reasonable to expect at least one such event $10^3 t_{\text{kyr}}$ yr ago, with $t_{\text{age}} \approx t_{\text{kyr}} \approx 2.0$. Then, for the current location of the SGR still to be at the centre of the ring to a precision of $1''$, the proper motion of the magnetar (due to its birth kick velocity v_{\perp}) and the proper motion of the dust-free cavity (with velocity v_c) should satisfy $\max(v_{\perp}, v_c) \leq 71 d_{15} t_{\text{kyr}}^{-1} \text{ km s}^{-1}$.

In light of the probable association between the SGR and the star cluster, we have re-examined the issue of the mismatch in extinction, which has been used as an argument against that connection¹¹. As our observations show, the cluster–SGR environment is characterized by a complex dust distribution on small spatial scales and dust destruction by the high-energy emission of the SGR. Because the derived extinction values were based on measurements at two distinct physical locations (the supergiants and the SGR) and are based on different tracers (gas and dust), it is not surprising that the resulting values differ. Similar differences have been observed for the Cas A SNR²⁹ and extragalactic gamma-ray-burst sources³⁰. In addition, we note that the visual absorption derived solely from the near-infrared photometry of the M supergiants, using both the published data¹² and new measurements extracted from the Two Micron All Sky Survey, is $A_V = 10.4\text{--}14.4$ mag, in comparison with $A_V = 19.2 \pm 1$ mag, the value obtained when the optical I-band measurement is included. Although this is puzzling and deserves further investigation, the various extinction measurements are clearly subject to large uncertainties and can potentially be reconciled with our conclusion that the SGR is a member of the star cluster.

Received 16 January; accepted 1 April 2008.

1. Duncan, R. C. & Thompson, C. Formation of very strongly magnetized neutron stars: implications for gamma-ray bursts. *Astrophys. J.* **392**, L9–L13 (1992).
2. Woods, P. M. & Thompson, C. in *Compact Stellar X-ray Sources* (eds Lewin, W. H. G. & van der Klis, M.) 547–586 (Cambridge Univ. Press, Cambridge, UK, 2006).
3. Figer, D. F., Najarro, F., Geballe, T. R., Blum, R. D. & Kudritzki, R. P. Massive stars in the SGR 1806–20 cluster. *Astrophys. J.* **622**, L49–L52 (2005).
4. Gaensler, B. M. *et al.* A stellar wind bubble coincident with the anomalous X-ray pulsar 1E 1048.1–5937: are magnetars formed from massive progenitors? *Astrophys. J.* **620**, L95–L98 (2005).

5. Muno, M. P. *et al.* A neutron star with a massive progenitor in Westerlund 1. *Astrophys. J.* **636**, L41–L44 (2006).
6. Kouveliotou, C. *et al.* An X-ray pulsar with a superstrong magnetic field in the soft γ -ray repeater SGR 1806–20. *Nature* **393**, 235–237 (1998).
7. Gaensler, B. M., Slane, P. O., Gotthelf, E. V. & Vasisht, G. Anomalous X-ray pulsars and soft gamma-ray repeaters in supernova remnants. *Astrophys. J.* **559**, 963–972 (2001).
8. Eikenberry, S. S. *et al.* Possible infrared counterparts to the soft gamma-ray repeater SGR 1806–20. *Astrophys. J.* **563**, L133–L137 (2001).
9. Heger, A., Fryer, C. L., Woosley, S. E., Langer, N. & Hartmann, D. H. How massive single stars end their life. *Astrophys. J.* **591**, 288–300 (2003).
10. Frail, D. A., Kulkarni, S. R. & Bloom, J. S. An outburst of relativistic particles from the soft gamma-ray repeater SGR 1900+14. *Nature* **398**, 127–129 (1999).
11. Kaplan, D. L., Kulkarni, S. R., Frail, D. A. & van Kerkwijk, M. H. Deep radio, optical, and infrared observations of SGR 1900+14. *Astrophys. J.* **566**, 378–386 (2002).
12. Vrba, F. J. *et al.* The double infrared source toward the soft gamma-ray repeater SGR 1900+14. *Astrophys. J.* **468**, 225–230 (1996).
13. Vrba, F. J. *et al.* The discovery of an embedded cluster of high-mass stars near SGR 1900+14. *Astrophys. J.* **553**, L17–L20 (2000).
14. Hurley, K. *et al.* ASCA discovery of an x-ray pulsar in the error box of SGR 1900+14. *Astrophys. J.* **510**, L111–L114 (1999).
15. Draine, B. T. Interstellar dust grains. *Annu. Rev. Astron. Astrophys.* **41**, 241–289 (2003).
16. Morris, P. W. *et al.* Tentative discovery of a new supernova remnant in Cepheus: unveiling an elusive shell in the Spitzer Galactic First Look Survey. *Astrophys. J.* **640**, L179–L182 (2006).
17. Chan, K. W. & Onaka, T. A broad 22 micron emission feature in the Carina Nebula H II region. *Astrophys. J.* **533**, L33–L36 (2000).
18. Smith, N., Bally, J. & Walawender, J. And in the darkness bind them: equatorial rings, B[e] supergiants, and the waists of bipolar nebulae. *Astron. J.* **134**, 846–859 (2007).
19. Weis, K. On the structure and kinematics of nebulae around LBVs and LBV candidates in the LMC. *Astron. Astrophys.* **308**, 67–71 (2003).
20. Gruendl, R. A., Chu, Y.-H., Dunne, B. C. & Points, S. D. A morphological diagnostic for dynamical evolution of Wolf–Rayet bubbles. *Astron. J.* **120**, 2670–2678 (2000).
21. Crowther, P. A. Physical properties of Wolf–Rayet stars. *Annu. Rev. Astron. Astrophys.* **45**, 177–219 (2007).
22. Cox, A. N. *Allen's Astrophysical Quantities* 4th edn (ed. Cox, A. N.) Ch. 7.5 15.3.1 (Springer, New York, 2000).
23. Koo, D.-C. & Heiles, C. A fast expanding H I shell in W44: A preexisting wind-blown shell overtaken by a supernova remnant. *Astrophys. J.* **442**, L679–L684 (1995).
24. Wachter, S., Kouveliotou, C., Patel, S., Figer, D. & Woods, P. Spitzer space telescope observations of SGR and AXP environments. *Astrophys. Space Sci.* **308**, 66–71 (2007).
25. Waxman, E. & Draine, B. T. Dust sublimation by gamma-ray bursts and its implications. *Astrophys. J.* **537**, 796–802 (2000).
26. Fruchter, A., Krolik, J. H. & Rhoads, J. E. X-ray destruction of dust along the line of sight to gamma-ray bursts. *Astrophys. J.* **563**, 597–610 (2001).
27. Tanaka, Y. T. *et al.* Comparative study of the initial spikes of soft gamma-ray repeater giant flares in 1998 and 2004 observed with Geotail: do magnetospheric instabilities trigger largescale fracturing of a magnetar's crust? *Astrophys. J.* **665**, L55–L58 (2007).
28. Palmer, D. A. *et al.* A giant gamma-ray flare from the magnetar SGR 1806–20. *Nature* **434**, 1107–1109 (2005).
29. Hartmann, D. H. *et al.* On Flamsteed's supernova Cas A. *Nucl. Phys. A* **621**, 83–91 (1997).
30. Stratta, G. *et al.* Extinction properties of the X-ray bright/optically faint afterglow of GRB 020405. *Astron. Astrophys.* **441**, 83–88 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work is based on observations made with the Spitzer Space Telescope, which is operated by the Jet Propulsion Laboratory (JPL), California Institute of Technology (Caltech), under a contract with NASA. Support for this work was provided by NASA through an award issued by JPL/Caltech. This publication also makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/Caltech, funded by NASA and the NSF. J.G. gratefully acknowledges a Royal Society Wolfson Research Merit Award. V.V.D. acknowledges support from the NSF and discussions with R. McCray, A. Crotts and R. Chevalier. D.F. acknowledges support from NASA through the Long Term Space Astrophysics programme, and by the New York State Foundation for Science, Technology, and Innovation Faculty Development Program grant.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.W. (wachter@ipac.caltech.edu).

A phase diagram for jammed matter

Chaoming Song¹, Ping Wang¹ & Hernán A. Makse^{1,2}

The problem of finding the most efficient way to pack spheres has a long history, dating back to the crystalline arrays conjectured¹ by Kepler and the random geometries explored² by Bernal. Apart from its mathematical interest, the problem has practical relevance³ in a wide range of fields, from granular processing to fruit packing. There are currently numerous experiments showing that the loosest way to pack spheres (random loose packing) gives a density of ~55 per cent^{4–6}. On the other hand, the most compact way to pack spheres (random close packing) results in a maximum density of ~64 per cent^{2,4,6}. Although these values seem to be robust, there is as yet no physical interpretation for them. Here we present a statistical description of jammed states⁷ in which random close packing can be interpreted as the ground state of the ensemble of jammed matter. Our approach demonstrates that random packings of hard spheres in three dimensions cannot exceed a density limit of ~63.4 per cent. We construct a phase diagram that provides a unified view of the hard-sphere packing problem and illuminates various data, including the random-loose-packed state.

Difficulties in describing static granular materials and other jammed systems, such as compressed emulsions, stem from the lack of well-defined conservation laws on which a statistical description of the system can be based. Unlike in equilibrium statistical mechanics, energy no longer describes the microstates of the system, owing to the dissipative and athermal nature of jammed matter. Thus, many experimental and theoretical studies focus on the analysis of the system volume as the analogue of system energy in equilibrium thermal systems^{7–16}. Recent advances in X-ray tomography¹³ and confocal microscopy¹⁷ have revealed the detailed internal structure of jammed matter, allowing for the study of the free volume per particle, or free volume function, denoted W (ref. 7). By partitioning the space into a set of non-overlapping volumes with Voronoi diagrams, these studies show that W is distributed with exponential tails^{13,14,17}. More importantly, experiments with monodisperse hard spheres¹³ show that W is inversely proportional to the coordination number (number of contacts) of the particle, z .

From a theoretical perspective, the study of the ensemble of jammed matter requires an analytical form for W (refs 7, 9, 10, 15). We first derive the Voronoi volume in terms of the particle positions (see Supplementary Information section IA) and then use statistical analysis to coarse-grain the Voronoi volume over a mesoscopic length scale, obtaining a mesoscopic free volume function (see Supplementary Information section IB) that is analytically tractable. For monodisperse hard spheres (grains) of volume V_g , we find:

$$W(z) = \frac{2\sqrt{3}}{z} V_g \quad (1)$$

The inverse relation with z is in general agreement with experiments¹³. The calculation of $W(z)$ is based on the environment of the grains, where each particle is assumed to be in a uniform background field produced by the other particles and not influenced by the

particle. Thus, equation (1) is akin to a quasi-particle theory: the coordination number z in equation (1) can be considered a coarse-grained average associated with ‘quasi-particles’ with free volume W . The key result is the relation between the Voronoi volume and the coordination number. This makes it possible to incorporate the volume function into a statistical mechanical description in terms of jammed hard spheres, using the constraint of mechanical stability as we show below.

The canonical partition function in the volume ensemble is the starting point for the statistical mechanics of jamming⁷, where the role traditionally played by the energy in thermal systems is replaced by the volume:

$$Q(X) = \int g(W) e^{-W/X} \Theta_W dW \quad (2)$$

Here X is the compactivity in units of volume⁷, determining the macrostates of the system (see below for an interpretation of this temperature-like parameter); $g(W)$ is the density of jammed states for a given volume W ; and Θ_W formally imposes the condition of jamming on the ensemble through force and torque balance. The main components of the theory are the uniformity assumptions behind the calculation of the mesoscopic volume function (see Supplementary Information section IC), the identification of the isostatic condition with the ensemble of jammed configurations, and the derivation of the density of states, as we discuss below.

Distinguishing between metastable and mechanically stable packings that define the jammed state through the Θ_W function is an unsolved problem¹⁸, and is related to the more fundamental question of whether or not a jammed packing is well defined. In practice, it is widely believed that the isostatic condition is necessary for a jammed disordered packing^{19–23}. That is, the number of force variables in the system is equal to the number of force and torque balance equations (see Supplementary Information section II). Therefore, we assume that the Θ_W function in equation (2) restricts the ensemble to the isostatic packings.

It is important to note that the coordination number z , as defined in equation (1), refers to the geometry of the packing and does not refer to contact forces. Therefore, z can include ‘trivial’ contacts with zero force, not contributing to the mechanical balance. We call z the geometrical coordination number to distinguish it from the mechanical coordination number Z , which is less than or equal to z and includes the contacts with non-zero force (see Supplementary Information section III). The mechanical coordination number thus directly corresponds to the isostatic condition of force and torque balance. For frictionless spherical particles (with interparticle friction coefficient $\mu = 0$, mimicking emulsion systems), the isostatic condition implies that $Z = 2d = 6$ (d is the dimension of the system, and in the following $d = 3$). For infinitely rough particles with $\mu \rightarrow \infty$, the mechanical coordination number is $Z = d + 1 = 4$ (see Supplementary Information section II for details). Interpolating between these two limits there must exist granular packings of finite

¹Levich Institute and Physics Department, City College of New York, New York, New York 10031, USA. ²Departamento de Física, Universidade Federal do Ceará, 60451-970 Fortaleza, Ceará, Brazil.

μ with $Z(\mu)$ smoothly varying between $Z(\mu = 0) = 6$ and $Z(\mu \rightarrow \infty) = 4$ (ref. 23). This is an important assumption that we test by numerical simulation (see Supplementary Information section II), where we find a common $Z(\mu)$ curve (Supplementary Fig. 10) for different packing preparation protocols. The mechanical coordination number ranges from four to six as a function of μ , and provides a lower bound on the geometrical coordination number: $Z \leq z \leq 6$. These bounds are tested in computer simulations in Supplementary Information section IIIA.

By changing variables, we can write equation (2) as (see Supplementary Information section IV):

$$Q_{\text{iso}}(X, Z) = \int_Z^6 e^{-W(z)/X} g(z) dz \quad (3)$$

Owing to the implicit volume coarse-graining in equation (1), each volume state $W(z)$ represents a mesoscopic state containing many microstates with a common value of z and density of states $g(z)$. The latter can be calculated as follows (see Supplementary Information section IV). We assume that the hard spheres are packed in a collectively jammed configuration in which no motion of any subset of particles can lead to unjamming²⁴. Thus, the configuration space of jammed matter is discrete, as we cannot continuously change one configuration to another. We denote the dimension per particle of the configuration space by \mathcal{D} and assume that the distance between two configurations is not broadly distributed, with a mean distance h_z . Therefore, the number of configurations is proportional to $1/(h_z)^{\mathcal{D}}$, analogous with that in quantum mechanics, h^{-d} , where h is Planck's constant and d is the dimension. The fact that the particles are jammed by z contacting particles reduces the number of degrees of freedom to $\mathcal{D} - z$, and the number of configurations is then $1/(h_z)^{\mathcal{D}-z}$. Because the term $1/(h_z)^{\mathcal{D}}$ is a constant, it will not influence the average in the partition function. Therefore, we have $g(z) = (h_z)^z$.

From equation (3) we obtain the equations of state that define the phase diagram of jamming. We start by investigating two limiting cases (see Supplementary Information section V). First, in the limit of vanishing compactivity ($X \rightarrow 0$), we obtain the ground state of jammed matter with a density

$$\phi_{\text{RCP}} = \frac{6}{6 + 2\sqrt{3}} \approx 0.634 \quad (4)$$

for $Z(\mu) \in [4, 6]$. Second, in the limit of infinite compactivity ($X \rightarrow \infty$), we obtain

$$\begin{aligned} \phi_{\text{RLP}}(Z) &= \frac{1}{Q_{\text{iso}}(\infty, Z)} \int_Z^6 \frac{z}{z + 2\sqrt{3}} (h_z)^z dz \\ &\approx \frac{Z}{Z + 2\sqrt{3}} \end{aligned} \quad (5)$$

for $Z(\mu) \in [4, 6]$.

The average in equation (5) is taken over all states with equal probability, because $e^{-W(z)/X} \rightarrow 1$ as $X \rightarrow \infty$, and the approximation applies because h_z is very small and the most populated state, $z = Z$, thus makes the dominant contribution to the average volume. The meaning of the subscripts 'RCP' (random close packing) and 'RLP' (random loose packing) in equations (4) and (5) will become clear below.

The equations of state (4) and (5) are plotted in the ϕ - Z plane in Fig. 1, illustrating the phase diagram of jammed matter. The phase space is limited to lie above the line of minimum coordination number, $Z = 4$ (for infinitely rough grains), labelled 'granular line' in Fig. 1. All mechanically stable, disordered jammed packings lie within the confining limits of the phase diagram (Fig. 1, yellow zone), and are forbidden in the grey area. For example, a packing of frictional hard spheres with $Z = 5$ (corresponding to a granular material with interparticle friction coefficient $\mu \approx 0.2$, according to Supplementary Fig. 10) cannot be equilibrated at volume fractions below

$\phi < \phi_{\text{RLP}}(Z = 5) = 5/(5 + 2\sqrt{3}) = 0.591$ or above $\phi > \phi_{\text{RCP}} = 0.634$. Thus, these results provide a statistical interpretation of the RLP and RCP limits, as follows.

First, originating in the statistical mechanics approach, the RCP limit arises as the result of equation (4), which gives the maximum volume fraction of disordered packings. The RCP density for monodisperse hard spheres^{2,4,6} is commonly quoted to be 63–64%; here we physically interpret a state with this value as the ground state of frictional hard spheres characterized by a given interparticle friction coefficient. In this representation, as μ varies from zero to infinity, the RCP state changes accordingly. This approach leads to an unexpected number of states lying in an 'RCP line' from the frictionless point at $Z = 6$ to the point at $Z = 4$, as depicted in Fig. 1, demonstrating that RCP is not a unique point in the phase diagram.

Second, equation of state (5) provides the lowest volume fraction for a given Z and represents a statistical interpretation of the RLP limit depicted by the 'RLP line' in Fig. 1. We predict that to the left of this line packings either are not mechanically stable or are experimentally irreversible as discussed in refs 8, 11, 25. There is no general consensus on the value of the RLP density: different estimates have been reported, ranging from 0.55 to 0.60 (refs 4–6). The phase diagram offers a solution to this problem. Along the infinite-compactivity RLP line, the volume fraction of the RLP decreases with increasing friction from the frictionless point $(\phi, Z) = (0.634, 6)$ (ref. 21), called the 'J-point' in ref. 22, towards the limit of infinitely rough hard spheres. Indeed, experiments⁴ indicate that lower volume fractions are associated with larger coefficients of friction. We predict the lowest volume fraction to be $\phi_{\text{RLP}}^{\text{min}} = 4/(4 + 2\sqrt{3}) \approx 0.536$, in the limit as $\mu \rightarrow \infty$, $X \rightarrow \infty$ and $Z \rightarrow 4$ ($h_z \ll 1$). Although this is a theoretical limit, our results indicate that for $\mu > 1$ this limit can be approximately achieved. The existence of an RLP bound is an interesting prediction of the present theory. The RLP limit has been little investigated experimentally, and currently it is not known whether this limit can be reached in real systems. Our prediction is close to the lowest stable volume fraction ever reported for monodisperse spheres⁵, namely 0.550 ± 0.006 .

Third, between the two RLP and RCP limits, there are packings inside the yellow zone in Fig. 1 with finite compactivity, $0 < X < \infty$. In such cases we solve the partition function numerically to obtain $\phi(X, Z)$ along an isocompactivity line, as shown in the colour lines in Fig. 1. The compactivity X controls the probability of each state, through a Boltzmann-like factor in equation (3) (as in condensed matter physics), and characterizes the number of possible ways of rearranging a packing having a given volume and entropy, S . Thus, the limits of the most compact and least compact stable arrangements correspond to $X \rightarrow 0$ and $X \rightarrow \infty$, respectively. Between these limits, the compactivity determines the volume fraction from RCP to RLP.

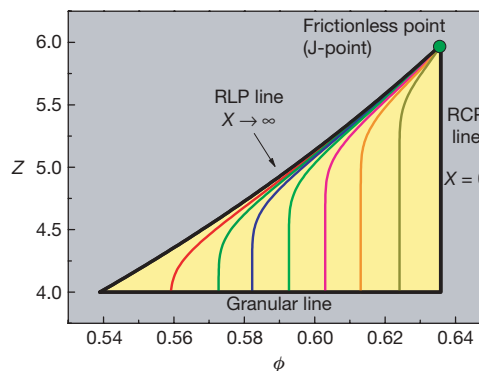


Figure 1 | Phase diagram of jamming: theory. Theoretical prediction of the statistical theory. All disordered packings lie within the yellow triangle demarcated by the RCP line, RLP line and granular line. Lines of uniform finite compactivity are in colour. Packings are forbidden in the grey area.

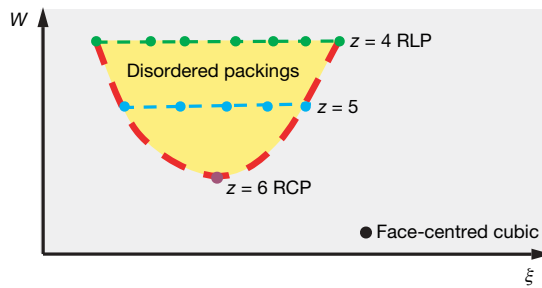


Figure 2 | Representation of the volume landscape of jammed matter (ξ , W). The multidimensional coordinate ξ represents the degrees of freedom: particle positions and rotations. Each dot represents a discrete jammed state at a given z ; those lying on a line of constant volume W share a common z value. We present the $\mu = \infty$ case. The states represent those along the granular line in Fig. 1 as the compactivity varies from $X = 0$ (ground state) to $X \rightarrow \infty$ (RLP limit). The ground state of jammed matter for this friction coefficient has $z = 6$ and the highest volume states are found for $z = 4$. For other finite values of μ , the space is delimited above by a line of constant $z = Z(\mu)$. All disordered packings lie in the yellow region of the phase space, which corresponds to the isostatic plane of hard spheres at the jamming transition where our calculations are performed. Other ordered packings, such as the face-centred cubic, have lower volume.

These results can be visualized in terms of a ‘volume landscape’ analogous to the energy landscape in glasses²⁶. Each jammed state (determined by the positions and rotations of the particles, denoted ξ , and the corresponding free volume, W) is depicted as a point in Fig. 2. The volume landscape has different levels of constant W determined by z , analogous to energy levels in hamiltonian systems. The lowest volume corresponds to the face-centred-cubic/hexagonal-close-packed structure (with $z = 12$), as conjectured by Kepler. Other lattice packings, such as the cubic lattice and tetrahedron lattice, have higher volume levels in this representation. Beyond these ordered states, the ensemble of disordered packings is identified within the yellow area in Fig. 2, corresponding to a system with infinite friction. Equation (4) indicates that the RCP corresponds to the ground state of disordered jammed matter for a given friction, which determines Z , whereas the RLP states are achieved for higher volume levels, as indicated in Fig. 2.

Further statistical characterization of the jammed structures can be obtained by solving the equations of state in the three-dimensional X – ϕ – S space, as in Fig. 3 (see Supplementary Information section V).

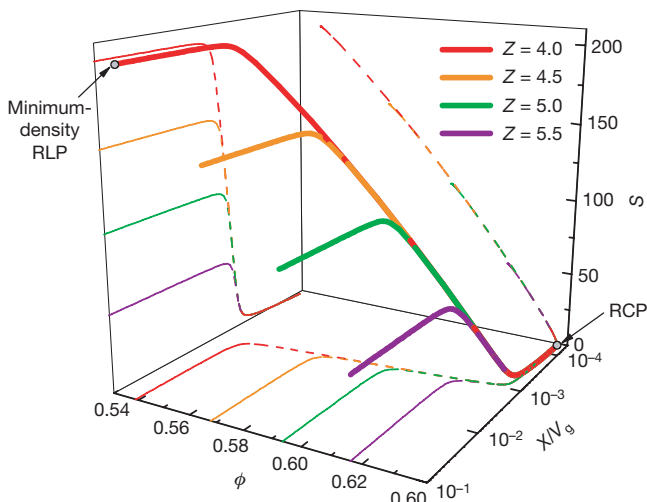


Figure 3 | Predictions of the equation of state of jammed matter in the X – ϕ – S space. Each line corresponds to a different system with Z as indicated. The projection in the X – ϕ plane qualitatively resembles the compaction curves of the experiments in refs 8, 11, 25.

Each curve in the figure corresponds to a system with a different $Z(\mu)$. The solution to the equation of state for $\phi(X)$ can be seen for different values of Z in the projection in Fig. 3. The volume fraction diminishes with increasing compactivity according to the theoretical picture of the phase diagram. The curves $\phi(X)$ qualitatively reproduce the reversible branch of compaction curves in the experiments of ref. 8 for shaken granular materials and oscillatory compression of grains²⁵, suggesting a correspondence between X and shaking amplitude. The idea is that different control parameters in experiments could be related to a state variable, and therefore might help experimentalists to describe results obtained under different protocols. For any value of Z , there is a common limit, $\phi \rightarrow \phi_{\text{RCP}}$, as $X \rightarrow 0$, giving the constant volume fraction for all the RCP states. The singular nature of the frictionless J-point is apparent from the fact that the volume fraction remains constant for any value of X , explaining why this point is the confluence of the isocompactivity lines, including RCP and RLP.

The existence of the theoretically inferred jammed states opens such predictions to experimental and computational investigation. We numerically test the predictions of the phase diagram by preparing monodisperse packings of Hertz–Mindlin spheres with friction coefficient μ at the jamming transition using previously developed methods^{21,27}. We obtain different packing states by compressing a system from an initial volume fraction ϕ_i with a compression rate Γ in a medium of viscosity (damping) η (see Supplementary Information section VIA). Although the simulations are not realistic (we do not use gravity, boundaries or a realistic protocol), they provide a way to test the main predictions of the theory. In Fig. 4 we plot the final state (ϕ, Z) reached by the system for every quadruplet $(\phi_i, \Gamma, \eta, \mu)$ at the jamming transition of vanishing stress, using a method explained in Supplementary Information section VIB. As in other non-equilibrium systems, such as glasses, the inherent path dependency of jammed matter is manifest in the fact that different packing structures can be realized using different preparation

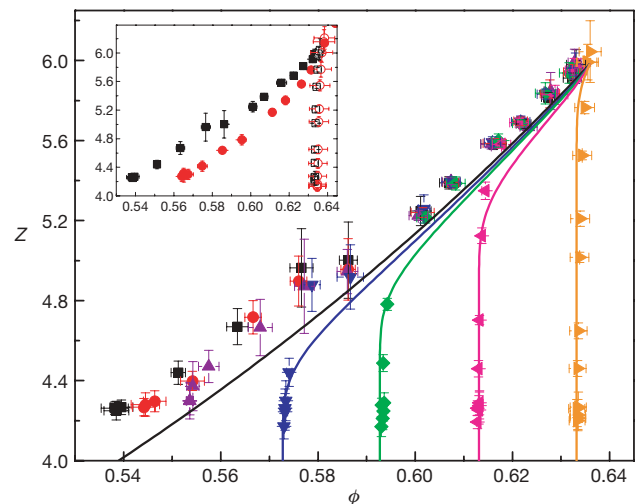


Figure 4 | Phase diagram of jamming: simulations. Numerical simulations demonstrate how to dynamically access the theoretically found states. The numerical protocol is parameterized by $(\phi_i, \Gamma, \eta, \mu)$. The main plot shows the dependence of the final jammed states (ϕ, Z) on ϕ_i for fixed $\Gamma = 10^{-7}$ and $\eta = 10^{-3}$ (except for the data plotted in orange, which is for $\eta = 10^{-4}$) and the ϕ_i values 0.40 (black), 0.53 (red), 0.55 (violet), 0.57 (blue), 0.59 (green), 0.61 (pink) and 0.63 (orange). For each ϕ_i , the different points have different values of μ (see Supplementary Fig. 10). Solid lines represent the theoretical results, with $h_z = e^{-100}$, for the following values of compactivity X (in units of $10^{-3}V_g$): infinity (black), 1.62 (blue), 1.38 (green), 1.16 (pink) and 0.88 (orange). The inset (which has the same axes as the main panel) focuses on the dependence of (ϕ, Z) on (Γ, η) for two different ϕ_i values. Filled symbols are for $\phi_i = 0.40$ and the (Γ, η) values $(10^{-7}, 10^{-3})$ (black) and $(10^{-4}, 10^{-4})$ (red). Open symbols are for $\phi_i = 0.63$ and the (Γ, η) values $(10^{-7}, 10^{-3})$ (black) and $(10^{-3}, 10^{-4})$ (red). The error bars correspond to the standard deviation over ten realizations of the packings.

protocols^{8,11,25}. Indeed, the present algorithm has analogies with recent attempts to describe jamming using ideas from the theory of mean-field spin glasses and optimization problems^{28,29}.

Changing the initial volume fraction ϕ_i produces different packings, as seen in Fig. 4. As ϕ_i increases, the final volume fraction approaches the prediction of the vertical RCP line of zero compactivity, demonstrating how to access the range of RCP states. All RCP states have approximately the same geometrical coordination number, $z \approx 6$, but differ in mechanical coordination number, with values ranging from $Z = 6$ to $Z \approx 4$, as predicted by the theory. When considering packings prepared with the smallest ϕ_i , slower compression rates (see Fig. 4 inset) or larger viscosities of the medium, we produce states with infinite compactivity along the RLP line. These results agree with the experiments of ref. 5: RLP is found for slowly deposited grains. Furthermore, packings prepared with intermediate values of ϕ_i closely follow the lines of isocompactivity shown in Fig. 4. Thus, to a reasonable approximation and for this particular protocol, we identify the density of the initial state, ϕ_i , with the compactivity of the packing, which provides a way to prepare a packing with a desired compactivity. In general, all numerically generated jammed states lie approximately within the predicted bounds of the phase diagram (see Supplementary Information section VIC for further details).

The numerical results indicate a way to test the existence of the predicted packings experimentally. By allowing the grains to settle in liquids of varying density, the speed of the particles can be varied and a systematic exploration of the jamming phase diagram can be made. Beyond the elucidation of some questions about the sphere-packing problem, other problems can now be addressed systematically using the phase diagram. These include the investigation of the criticality of the jamming transition from frictionless to frictional systems by extending the phase space to include stress; the characterization of jamming in the phase space of configurations; the problem of elasticity and Green's function; and the study of the distribution of forces, volumes and coordination numbers, to name a few. An advantage of the present formalism is that it provides a unified classification of jammed packings using which these studies could be systematically performed.

Received 3 December 2007; accepted 8 April 2008.

1. Hales, T. C. The Kepler conjecture. Preprint at (<http://arxiv.org/abs/math/9811078v2>) (2002).
2. Bernal, J. D. & Mason, J. Packing of spheres: co-ordination of randomly packed spheres. *Nature* **188**, 910–911 (1960).
3. Behringer, R. P. & Jenkins, J. T. (eds) *Powders & Grains* 97 (Balkema, Rotterdam, 1997).
4. Scott, G. D. & Kilgour, D. M. The density of random close packing of spheres. *J. Phys. D* **2**, 863–866 (1969).
5. Onoda, G. Y. & Liniger, E. G. Random loose packings of uniform spheres and the dilatancy effect. *Phys. Rev. Lett.* **64**, 2727–2730 (1990).
6. Berryman, J. D. Random close packing of hard spheres and disks. *Phys. Rev. A* **27**, 1053–1061 (1983).
7. Edwards, S. F. & Oakeshott, R. B. S. Theory of powders. *Physica A* **157**, 1080–1090 (1989).

8. Nowak, E. R., Knight, J. B., Ben-Naim, E., Jaeger, H. M. & Nagel, S. R. Density fluctuations in vibrated granular materials. *Phys. Rev. E* **57**, 1971–1982 (1998).
9. Blumenfeld, R. & Edwards, S. F. Granular entropy: explicit calculations for planar assemblies. *Phys. Rev. Lett.* **90**, 114303 (2002).
10. Ball, R. C. & Blumenfeld, R. Stress field in granular systems: loop forces and potential formulation. *Phys. Rev. Lett.* **88**, 115505 (2002).
11. Schröter, M., Goldman, D. I. & Swinney, H. L. Stationary state volume fluctuations in a granular medium. *Phys. Rev. E* **71**, 030301(R) (2005).
12. Fierro, A., Nicodemi, M., Tarzia, M., de Candia, A. & Coniglio, A. Jamming transition in granular media: A mean-field approximation and numerical simulations. *Phys. Rev. E* **71**, 061305 (2005).
13. Aste, T., Saadatfar, M. & Senden, T. J. Local and global relations between the number of contacts and density in monodisperse sphere packs. *J. Stat. Mech.* P07010 (2006).
14. da Cruz, F., Lechenault, F., Dauchot, O. & Bertin, E. Free volume distributions inside a bidimensional granular medium, in *Powders and Grains 2005* (eds García-Rojo, R., Herrmann, H. J. & McNamara, S.) (Balkema, Rotterdam, 2005).
15. Bertin, E., Dauchot, O. & Droz, M. Definition and relevance of nonequilibrium intensive thermodynamic parameters. *Phys. Rev. Lett.* **96**, 120601 (2006).
16. Ciamarra, M. P., Coniglio, A. & Nicodemi, M. Thermodynamics and statistical mechanics of dense granular media. *Phys. Rev. Lett.* **97**, 158001 (2006).
17. Bruijic, J., Edwards, S. F., Hopkinson, I. & Makse, H. A. Measuring distribution of interdroplet forces in a compressed emulsion system. *Physica A* **327**, 201–212 (2003).
18. Torquato, S., Truskett, T. M. & Debenedetti, P. G. Is random close packing of spheres well defined? *Phys. Rev. Lett.* **84**, 2064–2067 (2000).
19. Alexander, S. Amorphous solids: their structure, lattice dynamics and elasticity. *Phys. Rep.* **296**, 65–236 (1998).
20. Edwards, S. F. & Grinev, D. V. Statistical mechanics of stress transmission in disordered granular arrays. *Phys. Rev. Lett.* **82**, 5397–5400 (1999).
21. Makse, H. A., Johnson, D. L. & Schwartz, L. M. Packing of compressible granular materials. *Phys. Rev. Lett.* **84**, 4160–4163 (2000).
22. O'Hern, C. S., Langer, S. A., Liu, A. J. & Nagel, S. R. Random packings of frictionless particles. *Phys. Rev. Lett.* **88**, 075507 (2002).
23. Silbert, L. E., Ertas, D., Grest, G. S., Halsey, T. C. & Levine, D. Geometry of frictionless and frictional sphere packings. *Phys. Rev. E* **65**, 031304 (2002).
24. Torquato, S. & Stillinger, F. H. Multiplicity of generation, selection, and classification procedures for jammed hard-particle packings. *J. Phys. Chem. B* **105**, 11849–11853 (2001).
25. Bruijic, J. et al. Granular dynamics in compaction and stress relaxation. *Phys. Rev. Lett.* **95**, 128001 (2005).
26. Stillinger, F. H. A topographic view of supercooled liquids and glass formation. *Science* **267**, 1935–1939 (1995).
27. Zhang, H. P. & Makse, H. A. Jamming transition in emulsions and granular materials. *Phys. Rev. E* **72**, 011301 (2005).
28. Parisi, G. & Zamponi, F. The ideal glass transition of hard spheres. *J. Chem. Phys.* **123**, 144501 (2005).
29. Krzakala, F. & Kurchan, J. Landscape analysis of constraint satisfaction problems. *Phys. Rev. E* **76**, 021122 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work is supported by the National Science Foundation, CMMT Division and the US Department of Energy, Office of Basic Energy Sciences, Geosciences Division. We are grateful to J. Bruijic, A. Yupanqui and M. Makse for stimulating discussions.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to H.A.M. ([hmkse@lev.ccnycunyu.edu](mailto:hmakse@lev.ccnycunyu.edu)).

Quantum phase transition in a single-molecule quantum dot

Nicolas Roch¹, Serge Florens¹, Vincent Bouchiat¹, Wolfgang Wernsdorfer¹ & Franck Balestro¹

Quantum criticality is the intriguing possibility offered by the laws of quantum mechanics when the wave function of a many-particle physical system is forced to evolve continuously between two distinct, competing ground states¹. This phenomenon, often related to a zero-temperature magnetic phase transition, is believed to govern many of the fascinating properties of strongly correlated systems such as heavy-fermion compounds or high-temperature superconductors¹. In contrast to bulk materials with very complex electronic structures, artificial nanoscale devices could offer a new and simpler means of understanding quantum phase transitions^{2,3}. Here we demonstrate this possibility in a single-molecule quantum dot, where a gate voltage induces a crossing of two different types of electron spin state (singlet and triplet) at zero magnetic field. The quantum dot is operated in the Kondo regime, where the electron spin on the quantum dot is partially screened by metallic electrodes. This strong electronic coupling between the quantum dot and the metallic contacts provides the strong electron correlations necessary to observe quantum critical behaviour. The quantum magnetic phase transition between two different Kondo regimes is achieved by tuning gate voltages and is fundamentally different from previously observed Kondo transitions in semiconductor and nanotube quantum dots^{4,5}. Our work may offer new directions in terms of control and tunability for molecular spintronics⁶.

Quantum dots seem to be ideal devices in which to observe quantum phase transitions. First, such gate-tuneable artificial atoms offer a high degree of control by means of simple gate electrostatics. Second, owing to the nanometric confinement of the electrons, they display relatively high energy scales that allow the observation of interesting quantum phenomena at accessible temperatures. Finally, the coupling between the quantum dot and the electronic reservoirs (transport probes) produces tunnelling events that can fundamentally alter the discrete energy levels of the dot, changing them into complicated many-body wave functions. One well-studied situation in nanoscale devices (although not classified as a quantum transition) in which the interplay of these three effects is clear occurs when a single unpaired spin with $S = 1/2$ characterizes the ground state of the quantum dot. When conducting electrons move to and from the nanostructure, causing the tiny magnetic moment of the dot to alternate, a progressive screening of the atomic spin occurs, in complete analogy to the well-known Kondo effect in solids containing magnetic impurities^{7,8}. The Kondo effect in quantum dots is then observed as zero-bias conductance resonance^{9,10}, associated with the entangled state of electrons in the electronic reservoirs and in the dot, and displays a high degree of universality.

For quantum dots with even occupancy it is possible to tune the magnetic ground state of the dot between a spin-0 singlet and spin-1 triplet. Electronic tunnelling can then subtly affect the fate of the magnetic state of the quantum dot: when the singlet–triplet splitting

is eventually brought to zero, the simple level crossing of the two spin states becomes a true zero-temperature quantum phase transition. When only one electronic fluid participates in the Kondo effect^{2,3}, a spin-1 quantum dot will be partially compensated, so that the full many-body ground state evolves between different entropy states, presenting singularities at the quantum critical point (Fig. 1a). Conversely, the opening of a second screening channel^{11,12} will quench the remaining entropy, and the absence of obvious symmetry breaking generically leads to a transition being avoided^{13,14}.

Single-molecule quantum dots inserted in a nanoscale constriction present three features key to the observation of such sharp quantum phase transitions. First, owing to their asymmetric tunnelling geometry, a predominant single screening channel should be expected. Second, previous investigations have demonstrated large Kondo temperatures¹⁵. Third, as we discuss below, a gate-voltage dependence of the singlet–triplet gap for zero magnetic field can be introduced by means of a local gate¹⁵ underneath the nanogap, which allows precise tuning of the magnetic levels of the quantum dot. Combining these three crucial requirements in a single experiment opens new possibilities for the precise control of spin states in molecular nanostructures. Other than in the realization of interesting quantum many-body effects in quantum dots, our results may also be relevant in understanding quantum criticality in correlated bulk materials¹.

We used the electromigration technique¹⁶ (see Methods) to construct a single-molecule transistor, shown in Fig. 1b. Here we report on a full experimental study of transport measurements in terms of bias voltage V_b , gate voltage V_g , temperature T ($35 \text{ mK} < T < 20 \text{ K}$) and magnetic field B up to 8 T . Statistical evidence that transport takes place when we use a C_{60} molecule is provided in the Supplementary Information, together with several conductance characteristics obtained for different samples.

The general features of the single-molecule quantum dot are presented in Fig. 1c, which shows a large-scale, two-dimensional map of the differential conductance $\partial I / \partial V$ as a function of V_b and V_g at $T = 35 \text{ mK}$ and $B = 0$. The distinct conducting and non-conducting regions are typical signatures of a single-molecule transistor¹⁷. We present measurements over two distinct Coulomb diamonds indicated by ‘odd’ and ‘even’ charge states. The sharp high-conductance ridge in the odd charge state at zero bias is clearly associated with the usual spin-1/2 Kondo effect^{9,10}; detailed studies are reported in the Supplementary Information.

Henceforth we focus on the even charge state. The two-electron states can be described by their total spin S and spin projection m , and are denoted $|S, m\rangle$. The ground state of the system can thus be either a spin singlet $|0, 0\rangle$, with energy E_S , or a spin triplet, with energy E_T , described by the three states $|1, 1\rangle$, $|1, 0\rangle$ and $|1, -1\rangle$. These triplet states are degenerate for $B = 0$ but are split by the Zeeman effect, with an energy shift of $\Delta E_T = mg\mu_B B$ for each state $|1, m\rangle$, where g is the

¹Institut Néel, CNRS and Université Joseph Fourier, BP 166, 38042 Grenoble cedex 9, France.

g -factor and μ_B is the Bohr magneton. Figure 1d presents a precise low-bias $\partial I/\partial V$ map of the even region inside the dotted rectangle of Fig. 1c. This clearly displays two distinct regions, which (in anticipation of our results) we associate with the singlet and triplet ground states. The possibility of gate-tuning the singlet–triplet splitting $E_T - E_S$ was demonstrated previously both for lateral quantum dots¹⁸ and carbon nanotubes¹⁹, and may originate in an asymmetric coupling of the molecular levels to the electrodes²⁰. The magnetic states cross sharply at a critical gate voltage $V_g^c \approx 1.9$ V.

In the singlet region, a finite-bias conductance anomaly appears when V_b coincides with $E_T - E_S$; this is due to a non-equilibrium Kondo effect involving excitations into the spin-degenerate triplet. This effect was recently studied in a carbon-nanotube quantum dot in the singlet state²¹ (see Supplementary Information). In the triplet region, two kinds of resonance are observed: a finite-bias $\partial I/\partial V$ anomaly, which is interpreted as a singlet–triplet non-equilibrium Kondo effect that disperses like $E_S - E_T$ in the V_g – V_b plane, and a sharp, zero-bias $\partial I/\partial V$ peak, which is related to a partially screened spin-1 Kondo effect²², as indicated by the narrowness of the conductance peak.

To precisely identify these spin states, and justify our analysis in the framework of quantum criticality near the singlet–triplet crossing point, we present a detailed magneto-transport investigation of the even region. Owing to the high g -factor ($g \approx 2$) of C_{60} molecules, it is easy to lift the degeneracy of the triplet state for a C_{60} quantum dot using the Zeeman effect (see Supplementary Information). Figure 2b, d displays the evolution of the different conductance anomalies in the even region.

Figure 2b shows $\partial I/\partial V$ as a function of B and V_b for a constant gate voltage V_g chosen in the singlet region. A Zeeman-induced transition from the singlet state $|0, 0\rangle$ to the lowest- m triplet state $|1, -1\rangle$ occurs

as the magnetic field is increased (Fig. 2a), and is demonstrated by the clear level crossing in the conductance map. The splitting of the triplet is also apparent, and the various spectroscopic lines are consistent with the spin selection rules at both low and high magnetic field, where $|0, 0\rangle$ and $|1, -1\rangle$ are the respective ground states.

In Fig. 2d we investigate the gate-induced singlet–triplet crossing for constant magnetic field. In the singlet region, the Zeeman-split triplet states are clearly seen as three parallel lines, and the transition lines from the ground state $|1, -1\rangle$ at higher gate voltage are in agreement with the energy levels depicted in Fig. 2c, confirming the singlet to triplet crossing inside the even Coulomb diamond. We note the absence of a large enhancement of the zero-bias conductance at the singlet–triplet crossing in Figs 1d and 2b. Such features were, however, observed in previous experiments of vertical semiconductor quantum dots⁴, where a field-induced orbital effect can be used to make the non-degenerate triplet coincide with the singlet state, leading to a large Kondo enhancement of the conductance that is intimately related to the existence of two screening channels²³. In carbon nanotubes⁵, the Zeeman effect dominates over the orbital effect, so the transition involves the lowest- m triplet state only and Kondo signatures arise from a single channel, as in the case of well-balanced couplings of the two orbital states in the quantum dot to the electrodes²⁴. The lack of either type of singlet–triplet Kondo effect in our data indicates that the predominant coupling is between a single screening channel and one of the two spin states of the single-molecule quantum dot, leading to a Kosterlitz–Thouless quantum phase transition at the singlet–triplet crossing, as predicted by the theory^{2,3}. Although the peculiar magnetic response associated with this transition is not directly accessible in our scheme, we demonstrate that very specific characteristics of the Kosterlitz–Thouless transition can be observed in transport. The basic factor in the

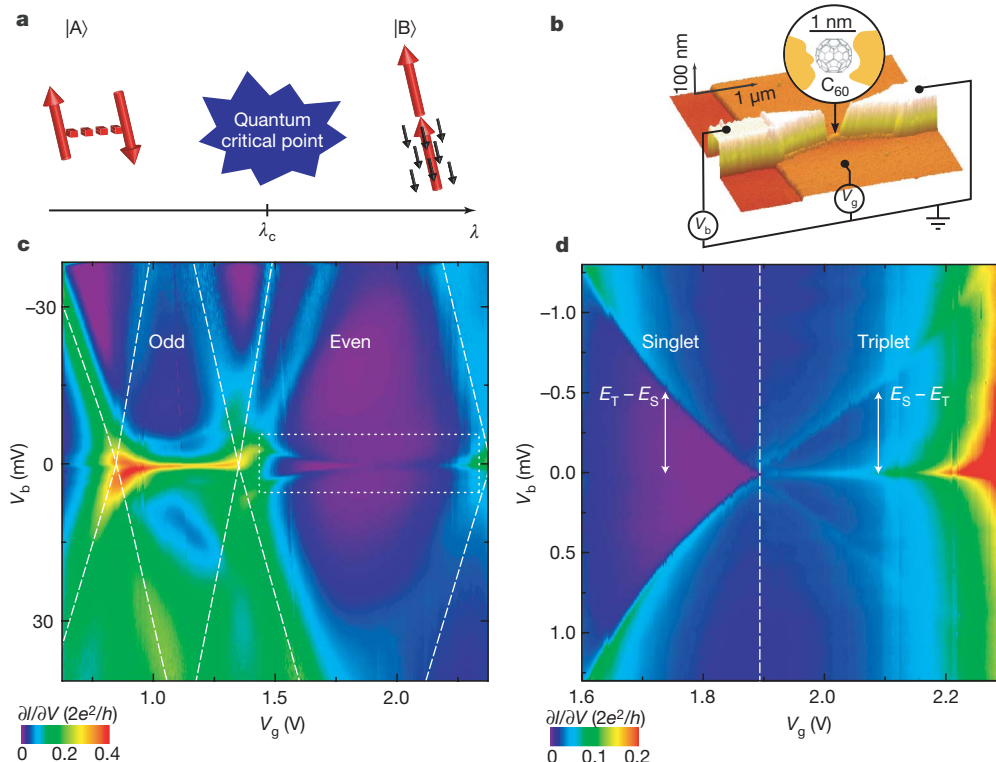


Figure 1 | Quantum phase transition, device and conductance characteristics. **a**, Quantum phase transition: a quantum state $|A\rangle$ can be driven by a non-thermal external parameter λ to another quantum state $|B\rangle$ with a different symmetry, passing through a critical point at $\lambda = \lambda_c$. In our single-molecule quantum dot device, $|A\rangle$ is a singlet state and $|B\rangle$ is a triplet state that is partially screened by one conduction electron channel, represented by black arrows. **b**, Atomic-force-microscope micrograph of the

device: gold nanowire over an $\text{Al}/\text{Al}_2\text{O}_3$ gate, with a C_{60} molecule trapped in the nanogap formed during the electromigration. **c**, Colour map over two Coulomb diamonds of the differential conductance $\partial I/\partial V$ (in units of $2e^2/h$, where h denotes Planck's constant) as a function of bias voltage V_b and gate voltage V_g at $T = 35$ mK and $B = 0$. **d**, Detail of the differential conductance in the dotted white rectangle in **c**, showing the singlet to triplet spin transition.

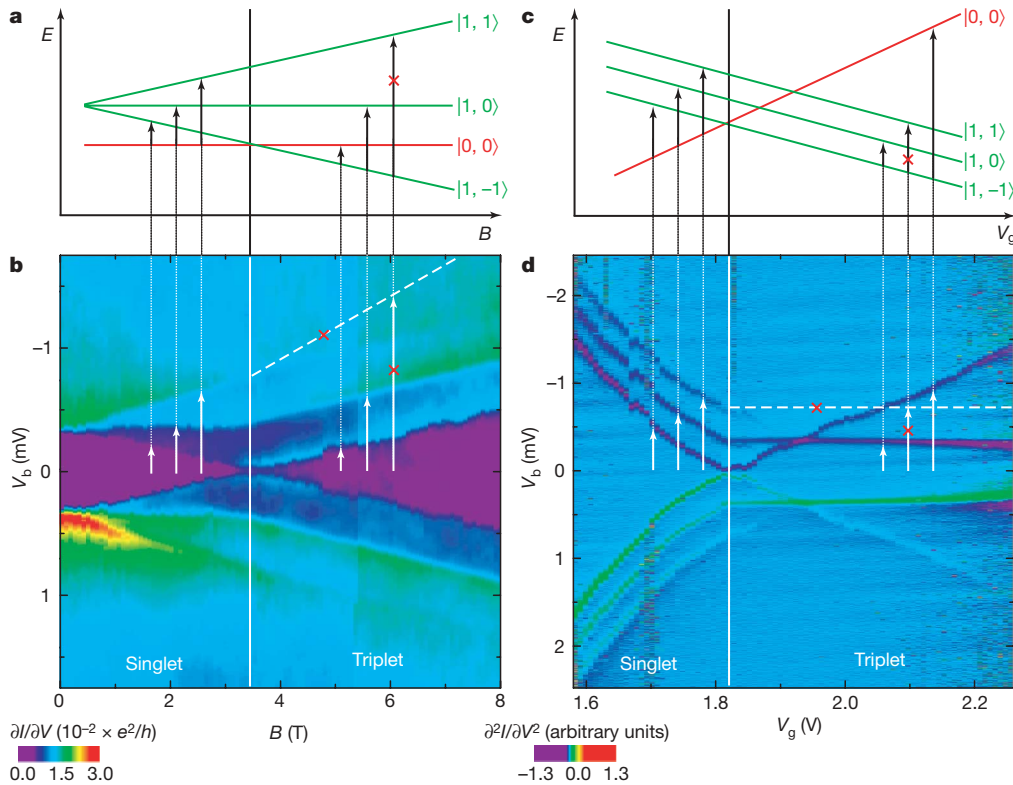


Figure 2 | Magnetic field and gate-induced singlet-triplet transition.

a, Transition from singlet state $|0, 0\rangle$ to lowest- m triplet state $|1, -1\rangle$ induced by the Zeeman effect. **b**, dI/dV measurements as a function of B and V_b at fixed gate voltage and temperature $T = 35$ mK. The crossed-out dashed line and arrow indicate that second-order spin-flip processes with $\Delta m = 2$ are not observed; see also **a**. **c**, Transition from singlet state $|0, 0\rangle$ to lowest- m triplet state $|1, -1\rangle$ induced by the gate voltage at constant magnetic field. **d**, d^2I/dV^2 measurements as a function of V_g and V_b for fixed magnetic field $B = 3$ T and temperature $T = 35$ mK. The crossed-out line and arrow indicate the same thing as in **b**; see also **c**. Owing to lower contrast in the triplet region, we plot d^2I/dV^2 here to improve the visibility of the data.

following discussion is that the singlet and triplet states fully dissociate into two independent spin-1/2 units near the Kosterlitz–Thouless transition, undergoing distinct Kondo screening processes.

Figure 3b shows the conductance map for gate voltages close to the critical value V_g^c , where singlet and triplet states are tuned to coincide for $B = 0$. We notice that a sharp conductance dip forms on the

singlet side of the transition, in contrast to the shallow minima observed in previous experiments for two-level quantum dots in the singlet regime^{18,19,25}. On the triplet side, we recognize the narrow zero-bias resonance of the spin-1 Kondo effect with a small Kondo temperature $T_{K,1}$. As the singlet–triplet splitting decreases on either side, the anomalous finite-bias features previously discussed

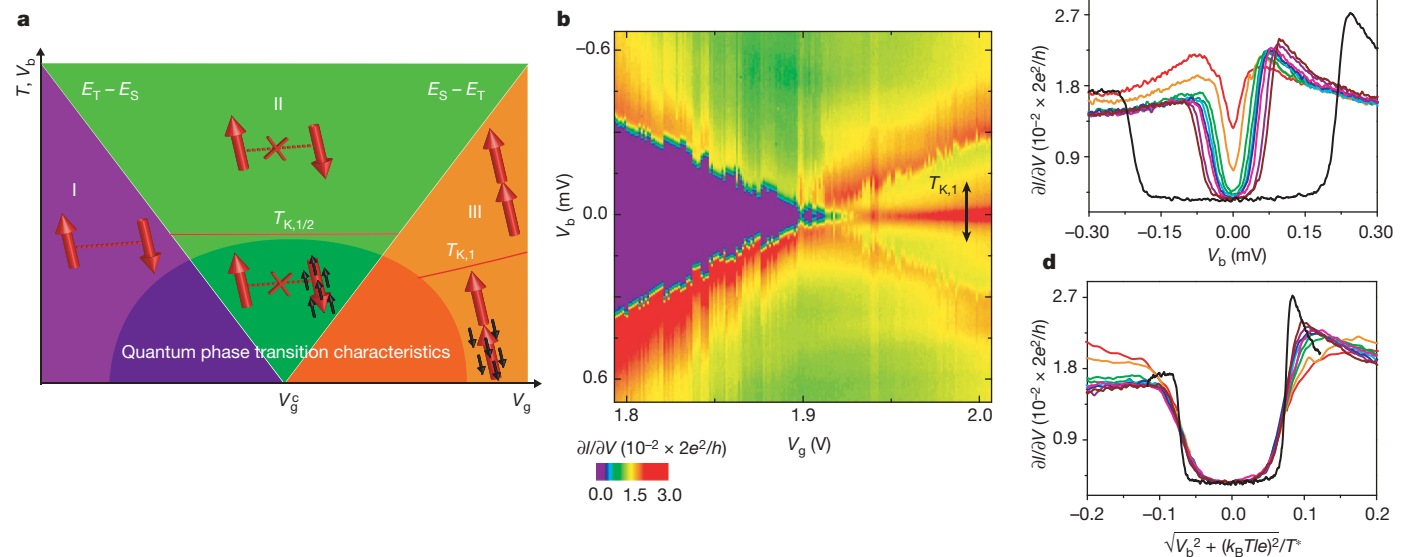


Figure 3 | Singlet-triplet quantum phase transition and universal scaling.

a, Phase diagram as a function of V_g and V_b (or T): we identify regions according to whether V_b (or T) lies above the singlet–triplet splitting $|E_T - E_S|$. The effective spin states of the quantum dot are represented by large red arrows and screening electrons are represented by smaller, black arrows. The dotted line between the spins in region I indicates a strongly bound singlet state. In region II the two spin states decouple from each other (indicated by a crossed-out dotted line), and the spin that is more strongly coupled to the leads is fully screened by means of a spin-1/2 Kondo effect

associated with the large Kondo temperature $T_{K,1/2}$. In region III the ground state of the quantum dot is a spin-1 triplet and experiences incomplete screening associated with the Kondo temperature $T_{K,1}$. **b**, Colour map of the differential conductance dI/dV as a function of V_b and V_g for $T = 35$ mK and $B = 0$ T, close to the singlet–triplet transition. **c**, Differential conductance for different values of $V_g < V_g^c$ close to (inverse Kondo effect exhibiting a resonant dip) and far from (U-shaped curve) the transition point. **d**, Scaling analysis of the data in **c**, with respect to the singlet binding energy $k_B T^*$.

smoothly merge to form a broad resonance related to a second, much larger, Kondo temperature $T_{K,1/2}$. Theoretical calculations² indicate that the physics of the critical point can be described in terms of one molecular level undergoing a standard spin-1/2 Kondo effect while the remaining spin-1/2 unit in the dot remains decoupled from the electrodes (as long as a second screening channel can be neglected).

The phase diagram in Fig. 3a depicts the proposed model of the singlet–triplet phase transition. By examining the temperature dependence of the conductance in the different regions shown in this plot, we can gain valuable insights into the interpretation in terms of a Kosterlitz–Thouless transition. In region I of Fig. 3a we identify two different regimes. Far from the transition point, that is, when the splitting $E_T - E_S$ exceeds the Kondo energy $k_B T_{K,1/2}$, where $T_{K,1/2}$ is the spin-1/2 Kondo temperature and k_B denotes Boltzmann's constant, the two spins strongly bind into an interorbital singlet. The differential conductance thus exhibits a characteristic 'U'-shape as a function of source–drain voltage (which is associated with the singlet–triplet gap), as shown by the wide curve in Fig. 3c. Close to the transition point, $E_T - E_S$ becomes less than $k_B T_{K,1/2}$, so Kondo screening of one of the orbital levels occurs before the freezing of the interorbital singlet. This results in both a broad resonance in the associated differential conductance (Fig. 4c) and a logarithmic increase in temperature of the zero-bias data (Fig. 4d). This logarithmic increase can be accurately fitted to the empirical formula²⁷ extracted from the numerical-renormalization-group calculations²⁷

$$G(T) = G_0 \left(\frac{T^2}{T_{K,1/2}^2} (2^{1/s} - 1) + 1 \right)^{-s} + G_c \quad (1)$$

where G_0 is the conductance at $T = 0$, G_c is a fixed background conductance and $s = 0.22$, giving $T_{K,1/2} = 4.13 \pm 0.3$ K.

Following further cooling, an interorbital singlet forms between the remaining unpaired spin and the first orbital level, which now is part of the Fermi sea due to its hybridization with the conduction electrons during the first stage of Kondo screening. The decrease in the zero-bias conductance is thus associated with a second-stage Kondo effect between these new degrees of freedom, as shown (Fig. 4c) by the formation below 600 mK of a characteristic narrow Kondo dip inside the broad resonance of width $T_{K,1/2}$. This interpretation is supported by the decrease in the zero-bias conductance (Fig. 4d), which is described in terms of an inverted Kondo peak, as the second-stage Kondo effect upon further cooling is associated with a re-entrant Kondo effect, by the formula

$$G(T) = G_0 \left(1 - \left(\frac{T^2}{T^{*2}} (2^{1/s} - 1) + 1 \right)^{-s} \right) + G_c \quad (2)$$

where now G_0 is a typical conductance value, giving the renormalized singlet binding energy $k_B T^*$ with $T^* = 187 \pm 21$ mK.

Because the formation of the singlet state close to the transition point is associated with a re-entrant Kondo effect, we should expect universal behaviour of the conductance dip with a characteristic temperature T^* . Figure 3c shows that, at the base temperature $T = 35$ mK and for $V_g < V_g^c$, the differential conductance evolves from a Kondo-like lorentzian shape to a U shape. These data are shown rescaled²⁸ as a function of $\sqrt{V_b^2 + (k_B T/e)^2}/T^*$, where e is the electron charge, in Fig. 3d. This plot shows that the conductance curves coincide when studied close to the transition point, but that the scaling deteriorates as the singlet–triplet gap $E_T - E_S$ becomes

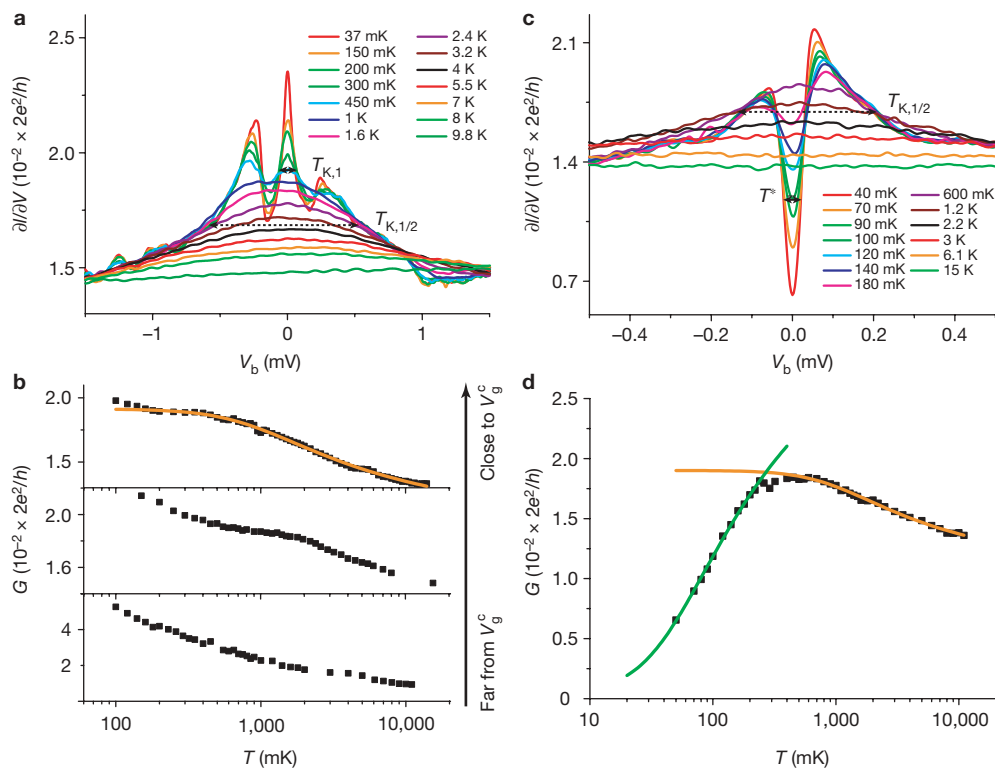


Figure 4 | Kondo effects in the singlet and triplet states. **a**, Differential conductance close to the transition point on the triplet side at different temperatures, showing a broad resonance with Kondo temperature $T_{K,1/2}$ and Kondo satellite peaks centred at $\pm |E_S - E_T|$. **b**, Temperature dependence of the zero-bias conductance $G(T)$ for three V_g values in the triplet region, the middle curve corresponding to **a**. We clearly measure a gate-dependent plateau, corresponding to the energy scale $|E_T - E_S|$. The solid line is a fit to equation (1), giving $T_{K,1/2} = 3.77 \pm 0.1$ K. The bottom panel, which does not show a clear plateau, corresponds to the temperature

evolution of $G(T)$ in the underscreened spin-1 regime. **c**, Differential conductance close to the transition point on the singlet side at different temperatures, showing a broad resonance with Kondo temperature $T_{K,1/2}$ and a narrow dip (inverse Kondo effect) associated with a temperature scale T^* . **d**, Temperature dependence of the zero-bias conductance $G(T)$ corresponding to **c**. The orange line is a fit to equation (1), giving $T_{K,1/2} = 4.13 \pm 0.3$ K, and the green line is a fit to equation (2), giving $T^* = 187 \pm 21$ mK.

greater than $k_B T_{K,1/2}$. Our study of the temperature dependence of $G(T)$ and the scaling of $\partial I/\partial V$ with V_g , both on the singlet side of the transition, thus provides strong evidence for a Kosterlitz–Thouless-like transition.

We now turn to region III in Fig. 3a, where the triplet state is favoured over the singlet state. Far from the transition point, at large values of V_g , the spins are tightly bound into a triplet and we expect to observe an underscreened spin-1 Kondo effect. Estimates made from both the width of the zero-bias peak and its magnetic field splitting (not shown) converge to a Kondo scale $T_{K,1}$ of the order of 100 mK. This value is too low to allow quantitative comparison with theoretical predictions of the underscreened Kondo effect, but the conductance data plotted in the bottom panel in Fig. 4b do not show any sign of saturation down to our effective electronic temperature $T_{\text{eff}} = 50$ mK.

For lower gate voltages, a complex regime in which the singlet–triplet splitting $E_S - E_T$ is comparable to the high-energy, single-level Kondo scale $T_{K,1/2}$ exists. This is shown, for fixed V_g and a range of temperatures, by the differential conductance data plotted in Fig. 4a. Although a broad peak is again observed at high temperatures, a three-peak structure emerges at low temperatures. We interpret the three-peak structure in terms of a non-equilibrium Kondo effect that mixes singlet and triplet states by means of the voltage-bias window. We associate the broad peak with a spin-1/2 Kondo effect, similar to that which occurs on the singlet side, in agreement with the corresponding zero-bias conductance $G(T)$ for temperatures above the singlet–triplet splitting temperature, giving $T_{K,1/2} = 3.77 \pm 0.1$ K (top panel in Fig. 4b). This spin-1/2 Kondo behaviour remains similar down to the lowest temperatures by approaching the critical point, as $E_S - E_T$ becomes smaller than $k_B T_{K,1/2}$. The further increase of $G(T)$ below $T = 200$ mK is at present not fully understood, and may be related to the opening of a second screening channel, which might spoil the quantum critical point at $T = 0$ K^{23,29}. However, this extra feature seems to be relevant only at a very small energy scale close to the crossing point, so the data in the accessible temperature range is consistent within the quantum critical point interpretation.

We finally note that our interpretation of the experimental data can be supported in a complementary way (see Supplementary Information), namely plotting the zero-bias conductance as a function of gate voltage for different temperatures. As the temperature decreases, we observe the clear sharpening of a conductance step when the system crosses from the singlet region to the triplet region, in agreement with the existence of a quantum critical point³ and in contrast to the maximum predicted for an avoided transition^{23,29}.

METHODS SUMMARY

The single-molecule transistors were produced using standard electron-beam lithography to pattern a gold nanowire deposited on an Al/Al₂O₃ back gate; see Fig. 1b for an atomic-force-microscope micrograph with a diagram of the set-up. The nanowire junctions were cleaned with acetone, ethanol, isopropanol solution and oxygen plasma. Then a dilute toluene solution of C₆₀ molecules was deposited on the junctions and blow-dried. The connected samples were placed in a copper shielded box that had high-frequency low-temperature filters. The box was anchored to the mixing chamber of the dilution fridge with a base temperature of 35 mK. The nanowire coated with molecules was then broken by electromigration¹⁶ at 4 K, using fast real-time electronics to control the coupling of the single molecule to the electrodes.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 December 2007; accepted 18 March 2008.

1. Sachdev, S. Quantum magnetism and criticality. *Nature Phys.* **4**, 173–185 (2008).

2. Vojta, M., Bulla, R. & Hofstetter, W. Quantum phase transitions in models of coupled magnetic impurities. *Phys. Rev. B* **65**, 140405 (2002).
3. Hofstetter, W. & Schoeller, H. Quantum phase transition in a multilevel dot. *Phys. Rev. Lett.* **88**, 016803 (2002).
4. Sasaki, S. *et al.* Kondo effect in an integer-spin quantum dot. *Nature* **405**, 764–767 (2000).
5. Nygård, J., Cobden, D. H. & Lindelof, P. E. Kondo physics in carbon nanotubes. *Nature* **408**, 342–346 (2000).
6. Bogani, L. & Wernsdorfer, W. Molecular spintronics using single-molecule magnets. *Nature Mater.* **7**, 179–186 (2008).
7. Hewson, A. C. *The Kondo Problem to Heavy Fermions* (Cambridge Univ. Press, Cambridge, UK, 1993).
8. Glazman, L. I. & Raikh, M. E. Resonant Kondo transparency of a barrier with quasilocal impurity states. *JETP Lett.* **47**, 452–455 (1988).
9. Goldhaber-Gordon, D. *et al.* Kondo effect in a single-electron transistor. *Nature* **391**, 156–159 (1998).
10. Cronenwett, S. M., Oosterkamp, T. H. & Kouwenhoven, L. P. A tunable Kondo effect in quantum dots. *Science* **281**, 540–544 (1998).
11. Georges, A. & Meir, Y. Electronic correlations in transport through coupled quantum dots. *Phys. Rev. Lett.* **82**, 3508–3511 (1999).
12. Jones, B. A., Varma, C. M. & Wilkins, J. W. Low-temperature properties of the two-impurity Kondo hamiltonian. *Phys. Rev. Lett.* **61**, 125–128 (1988).
13. Affleck, I., Ludwig, A. W. W. & Jones, B. A. Conformal-field-theory approach to the two-impurity Kondo problem: Comparison with numerical renormalization-group results. *Phys. Rev. B* **52**, 9528–9546 (1995).
14. Zarand, G., Chung, C.-H., Simon, P. & Vojta, M. Quantum criticality in a double-quantum-dot system. *Phys. Rev. Lett.* **97**, 166802 (2006).
15. Liang, W., Shores, M. P., Bockrath, M., Long, J. R. & Park, H. Kondo resonance in a single-molecule transistor. *Nature* **417**, 725–729 (2002).
16. Park, H., Lim, A. K. L., Alivisatos, A. P., Park, J. & McEuen, P. L. Fabrication of metallic electrodes with nanometer separation by electromigration. *Appl. Phys. Lett.* **75**, 301–303 (1999).
17. Park, H. *et al.* Nanomechanical oscillations in a single-C₆₀ transistor. *Nature* **407**, 57–60 (2000).
18. Kogan, A., Granger, G., Kastner, M. A., Goldhaber-Gordon, D. & Shtrikman, H. Singlet–triplet transition in a single-electron transistor at zero magnetic field. *Phys. Rev. B* **67**, 113309 (2003).
19. Quay, C. H. L. *et al.* Magnetic field dependence of the spin-1/2 and spin-1 Kondo effects in a quantum dot. *Phys. Rev. B* **76**, 245311 (2007).
20. Holm, J. V. *et al.* Gate-dependent tunneling-induced level shifts in carbon nanotube quantum dots. Preprint at (<http://aps.arxiv.org/abs/0711.4913>) (2007).
21. Paaske, J. *et al.* Non-equilibrium singlet–triplet Kondo effect in carbon nanotubes. *Nature Phys.* **2**, 460–464 (2006).
22. Nozières, P. & Blandin, A. Kondo effect in real metals. *J. Phys. (Paris)* **41**, 193–211 (1980).
23. Pustilnik, M. & Glazman, L. I. Kondo effect induced by a magnetic field. *Phys. Rev. B* **64**, 045328 (2001).
24. Pustilnik, M., Avishai, Y. & Kikoin, K. Quantum dot with even number of electrons: Kondo effect in a finite magnetic field. *Phys. Rev. Lett.* **84**, 1756–1759 (2000).
25. Craig, N. J. *et al.* Tunable nonlocal spin control in a coupled-quantum dot system. *Science* **304**, 565–567 (2004).
26. Grobis, M., Rau, I. G., Potok, R. M. & Goldhaber-Gordon, D. Kondo effect in mesoscopic quantum dots, in *Handbook of Magnetism and Magnetic Materials* Vol. 1 (eds Kronmüller, H. & Parkin, S.) Part II (Wiley, Chichester, 2007).
27. Costi, T. A. Kondo effect in a magnetic field and the magnetoresistivity of Kondo alloys. *Phys. Rev. Lett.* **85**, 1504–1507 (2000).
28. Potok, R. M., Rau, I. G., Shtrikman, H., Oreg, Y. & Goldhaber-Gordon, D. Observation of the two-channel Kondo effect. *Nature* **446**, 167–171 (2006).
29. Hofstetter, W. & Zarand, G. Singlet–triplet transition in lateral quantum dots: A numerical renormalization group study. *Phys. Rev. B* **69**, 235301 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We acknowledge E. Eyraud and D. Lepoittevin for discussions and technical contributions regarding electronics and dilutions. We thank E. Bonet, T. Crozes and T. Fournier for lithography development, and C. Winkelmann, T. Costi and L. Calvet for discussions. The sample used in the investigations was made in the NANOFAB facility of the Néel Institut. This work is partially financed by ANR-PNANO, Contract MolSpintronics.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to F.B. (franck.balestro@grenoble.cnrs.fr).

METHODS

The spin-1/2 Kondo effect in a C_{60} molecular junction was observed for the first time by Yu and Natelson³⁰ (see also the case of ferromagnetic electrodes³¹), and more recently by Parks *et al.*³² using mechanically controllable break junctions. Although improvements of the original electromigration procedure¹⁶ have also been reported recently^{33–39}, to our knowledge no electromigration has been carried out in a dilution refrigerator with a high degree of filtering. Because the creation of nanogaps using the electromigration technique requires minimizing the series resistance³⁹, the resistive dilution fridge wiring and filtering are generally incompatible with electromigration. However, accessing very low electronic temperatures is central to the observation of quantum critical signatures associated with the singlet–triplet crossing in the C_{60} molecular junction. To overcome this problem, we developed a specific electromigration set-up accessing very low electronic temperatures.

Our experimental method is divided into two parts (see Supplementary Fig. 1). In the first step, electromigration is performed at 4 K with the fast electronics of the set-up. We ramp the voltage across the junction and measure its resistance, using a very fast feedback loop (1.5 μ s) to set the voltage to zero when the resistance exceeds a defined threshold, typically 20 k Ω . The fast feedback was achieved using an ADwin-Pro II real-time data acquisition system and a home-built high-bandwidth current-to-voltage converter. With this technique, and a series resistance due to the filters of about 200 Ω , we obtained small gaps (1–2 nm), which we characterized in terms of the tunnel current measured after electromigration of junctions without molecules.

The second step uses the low-noise component of the set-up to measure the single-molecule transistor. In addition to low-temperature filtering using a thermocoax microwave filter and Π filters, we used Π filters and ferrite-bead filters of special design with a cutoff frequency of 100 MHz (the details of these filters can be found at http://marcuslab.harvard.edu/how_to/Ferrite_Bead_Filter.pdf). To minimise ground loops we integrated all the analogue electronics in a shielded box at room temperature. Because of its great versatility, the ADwin-Pro II system can be programmed to perform direct-current and lock-in measurements, and apply gate or bias voltages, thus minimizing the possibility of ground loops. Depending on the measurements, we used an alternating-current excitation between 3 μ V and 100 μ V for the lock-in technique.

We note that previous studies of C_{60} quantum dots did not require the use of a dilution refrigerator to investigate Kondo physics, because the relevant energy scales are typically an order of magnitude larger than in carbon nanotubes or semiconducting devices, providing large Kondo temperatures of several Kelvins. However, the study of single-molecule transistors using low-temperature techniques (previously reserved to two-dimensional electron gases) was crucial for observing the rich physics that takes place below the Kondo temperature at the singlet–triplet transition. Our low-temperature set-up also allowed a more precise investigation of the usual spin-1/2 Kondo effect in C_{60} (see Supplementary Information section 2).

30. Yu, L. H. & Natelson, D. The Kondo effect in C_{60} single-molecule transistors. *Nano Lett.* **4**, 79–83 (2004).
31. Pasupathy, A. N. *et al.* The Kondo effect in the presence of ferromagnetism. *Science* **306**, 86–89 (2004).
32. Parks, J. J. *et al.* Tuning the Kondo effect with a mechanically controllable break junction. *Phys. Rev. Lett.* **99**, 026601 (2007).
33. Strachan, D. R. *et al.* Controlled fabrication of nanogaps in ambient environment for molecular electronics. *Appl. Phys. Lett.* **86**, 043109 (2005).
34. Houck, A. A., Labaziewicz, J., Chan, E. K., Folk, J. A. & Chuang, I. L. Kondo effect in electromigrated gold break junctions. *Nano Lett.* **5**, 1685–1688 (2005).
35. Esen, G. & Fuhrer, M. S. Temperature control of electromigration to form gold nanogap junctions. *Appl. Phys. Lett.* **87**, 263101 (2005).
36. Trouwborst, M. L., van der Molen, S. J. & van Wees, B. J. The role of Joule heating in the formation of nanogaps by electromigration. *J. Appl. Phys.* **99**, 114316 (2006).
37. O'Neill, K., Osorio, E. A. & van der Zant, H. S. J. Self-breaking in planar few-atom Au constrictions for nanometer-spaced electrodes. *Appl. Phys. Lett.* **90**, 133109 (2007).
38. Wu, Z. M. *et al.* Feedback controlled electromigration in four-terminal nanojunctions. *Appl. Phys. Lett.* **91**, 053118 (2007).
39. van der Zant, H. S. J. *et al.* Molecular three-terminal devices: fabrication and measurements. *Faraday Discuss.* **131**, 347–356 (2006).

LETTERS

Anatase TiO₂ single crystals with a large percentage of reactive facets

Hua Gui Yang^{1*}, Cheng Hua Sun^{1,2*}, Shi Zhang Qiao¹, Jin Zou³, Gang Liu^{1,4}, Sean Campbell Smith^{1,2}, Hui Ming Cheng⁴ & Gao Qing Lu¹

Owing to their scientific and technological importance, inorganic single crystals with highly reactive surfaces have long been studied^{1–13}. Unfortunately, surfaces with high reactivity usually diminish rapidly during the crystal growth process as a result of the minimization of surface energy. A typical example is titanium dioxide (TiO₂), which has promising energy and environmental applications^{14–17}. Most available anatase TiO₂ crystals are dominated by the thermodynamically stable {101} facets (more than 94 per cent, according to the Wulff construction¹⁰), rather than the much more reactive {001} facets^{8–13,18–20}. Here we demonstrate that for fluorine-terminated surfaces this relative stability is reversed: {001} is energetically preferable to {101}. We explored this effect systematically for a range of non-metallic adsorbate atoms by first-principle quantum chemical calculations. On the basis of theoretical predictions, we have synthesized uniform anatase TiO₂ single crystals with a high percentage (47 per cent) of {001} facets using hydrofluoric acid as a morphology controlling agent. Moreover, the fluorinated surface of anatase single crystals can easily be cleaned using heat treatment to render a fluorine-free surface without altering the crystal structure and morphology.

The surface stability and reactivity of inorganic single crystals have long been thought to be dominated by their surface chemistry, whose effect on the equilibrium morphology is critical for the synthesis of single crystals with high reactivity^{1–13,21}. For anatase TiO₂, both theoretical and experimental studies found that the minority {001} facets in the equilibrium state are especially reactive⁸. However, large high-quality anatase single crystals with a high percentage of {001} facets have not been synthesized^{22–24}. An early study²³ showed that the hydrothermal treatment of hydrous titanium(IV) oxide in the presence of hydrofluoric acid resulted in irregular aggregates of polymorphic TiO₂ with anhedral morphology. Recently, anatase single crystals were synthesized using chemical transport reactions, but the process had a long reaction time and the crystals were of low purity and had no {001} facets²⁴. Therefore, preparation of uniform, high-purity anatase single crystals with controllable crystallographic facets still remains a challenge.

To this end, attempts have been made with various adsorbate atoms to change the relative stabilities of different crystal facets^{19–23}. For anatase TiO₂, among oxygenated surfaces, {100} facets are the most stable, whereas under clean and hydrogenated conditions, {101} facets are the most stable^{5,10,21}. However, both H- and O-terminated anatase surfaces present high surface energies (γ), which restrict the formation of large anatase single crystals. High values of γ are mainly attributed to the high bonding energies (D_0) of H–H (436.0 kJ mol^{–1}) and O–O (498.4 kJ mol^{–1})²⁵. Therefore, using a low- D_0 element with strong bonding to Ti might provide an

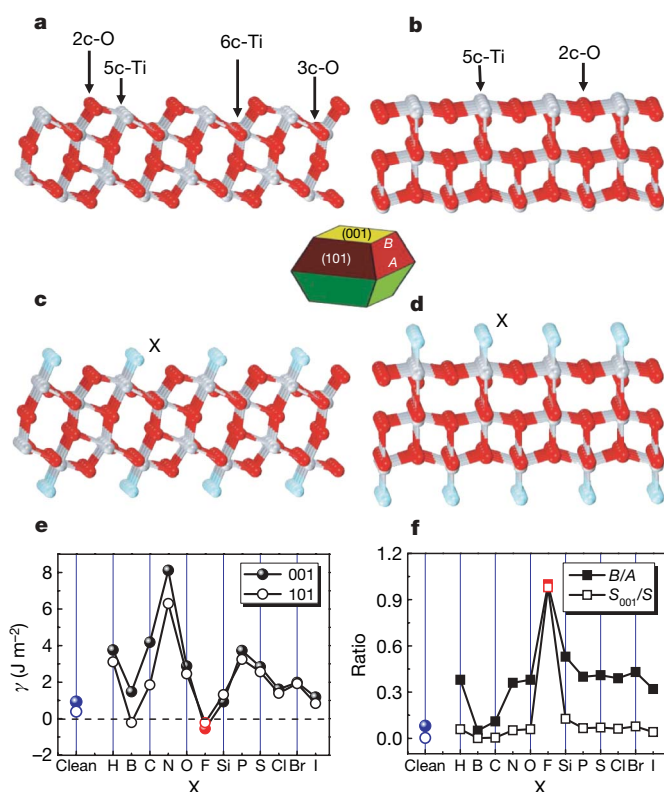


Figure 1 | Slab models and calculated surface energies of anatase TiO₂ {001} and {101} surfaces. The optimised ratios of B/A and percentage of {001} facets (S_{001}/S), where S and S_{001} are respectively the total surface area and that contributed by the {001} facets, are also shown. **a, b**, Unrelaxed, clean {001} and {101} surfaces. Ti and O atoms are represented by grey and red spheres, with sixfold Ti, fivefold Ti, threefold O and twofold O labelled as 6c-Ti, 5c-Ti, 3c-O and 2c-O, respectively. **c, d**, Unrelaxed {001} and {101} surfaces surrounded by adsorbate X atoms. **e**, Calculated energies of the {001} and {101} surfaces surrounded by X atoms. **f**, Plots of the optimized value of B/A and percentage of {001} facets for anatase single crystals with various adsorbate atoms X. In **e** and **f**, clean-surface results (denoted by blue spheres and circles) are used for reference. As indicated in the inset diagram, two independent parameters A and B denote lengths of the side of the bipyramid and the side of the square {001} ‘truncation’ facets, respectively. The ratio of highly reactive {001} facets to total surface area may therefore be described by the value of S_{001}/S or B/A (where $0 \leq B/A \leq 1$).

¹ARC Centre of Excellence for Functional Nanomaterials, School of Engineering and Australian Institute for Bioengineering and Nanotechnology, ²Centre for Computational Molecular Science, Australian Institute for Bioengineering and Nanotechnology, ³Centre for Microscopy and Microanalysis and School of Engineering, The University of Queensland, Queensland 4072, Australia. ⁴Shenyang National Laboratory for Materials Science, Institute of Metal Research, Chinese Academy of Sciences, 72 Wenhua Road, Shenyang 110016, China.

*These authors contributed equally to this work.

effective means for stabilizing the surfaces. Interestingly, F is such an element, as $D_0^{\text{F}-\text{F}} = 158.8 \text{ kJ mol}^{-1}$ (ref. 25) and $D_0^{\text{F}-\text{Ti}} = 569.0 \text{ kJ mol}^{-1}$ (ref. 26).

To further explore the effects of various adsorbate atoms, we carried out a systematic investigation of 12 non-metallic atoms X (where X can represent H, B, C, N, O, F, Si, P, S, Cl, Br or I) based on first-principle calculations. Figure 1a–d illustrates the models of clean and X-terminated (001) and (101) surfaces. The calculated γ values for different adsorbates are shown in Fig. 1e, from which two conclusions can be drawn: among the 12 X-terminated surfaces and the clean surfaces, termination with F atoms not only yields the lowest value of γ for both the (001) and (101) surfaces, but also results in (001) surfaces that are more stable than (101) surfaces.

These results indicate that it might be possible to achieve anatase TiO_2 single crystals with a high percentage of anatase {001} facets if their surfaces are surrounded by F atoms. Furthermore, on the basis of the shape-dependent thermodynamic model proposed in ref. 27, the optimized value of B/A (the length parameters illustrated in Fig. 1) and the percentage of {001} facets can be predicted if γ is known. Fig. 1f shows that the F-terminated surfaces have the highest degree of truncation ($B/A \rightarrow 1$ approximately); the F-terminated surfaces of anatase TiO_2 should therefore be dominated by {001} facets.

To verify these theoretical predictions, titanium tetrafluoride (TiF_4) aqueous solution and hydrofluoric acid were used as the anatase single crystals' precursor and crystallographic controlling agent, respectively, to generate the truncated anatase bipyramids. Representative scanning electron microscopy (SEM) images of the products synthesized with different concentrations of TiF_4 and reaction times are shown in Fig. 2a, d. On the basis of the symmetries of anatase TiO_2 , the two flat, square surfaces must be {001} facets and the eight isosceles trapezoidal surfaces are {101} facets of the anatase TiO_2 single crystals (further evidence for this is in Fig. 3 and Supplementary Fig. 7). The yield of anatase TiO_2 single crystals is around 90%, even though some agglomerates and/or irregular particles (Supplementary Fig. 4) were occasionally observed.

To examine the uniformity of the synthesized anatase crystals, we statistically analysed the values of A and B/A under the different synthesis conditions. The results are presented in Fig. 2b, c, e, f: for the respective cases shown in Fig. 2a, d, the average values of A are $1.66 \mu\text{m}$ and $1.64 \mu\text{m}$ with relative standard deviations of 8.4% and 15.8% (Fig. 2b, e), and the degrees of truncation (given by B/A) are 0.77 and 0.84 with relative standard deviations of 4.3% and 5.1% (Fig. 2, f). We estimate the percentages of {001} facets to be 35% and 47%, respectively.

The fact that anatase single crystals generated with a low concentration of TiF_4 exhibit a high degree of truncation may be attributed to the higher fluorine density on the surface, which makes the isotropic growth more obvious. This is remarkably consistent with our theoretical predictions, and can be understood from the viewpoint of shape-control chemistry^{20,21}. The free energy of (001) surfaces can be reduced to 0.51 J m^{-2} using the (1×4) reconstruction¹³, which suggests that this reconstruction could also stabilize (001) surfaces. However, on the basis of the values of γ calculated in ref. 13, the optimized value of B/A is around 0.37. This is very different from the values, 0.77 and 0.84, that we obtained in our experiments, indicating that the high percentage of {001} facets does not result from the (1×4) reconstruction. According to our detailed calculations based on a perfect (1×1) unit cell and (4×4) supercells of (001) surfaces (Supplementary Information Part I), all surfaces terminated with F atoms have very low values of γ . We therefore believe that the stabilization effect of F atoms is the essential reason for the percentage of {001} facets being as high as 47% in our experiments.

Regarding the stabilization mechanism, a detailed analysis of the surface geometries has been carried out (Supplementary Information Part I). Under clean conditions, the balance between the O–O repulsions and the attractive Ti–O π interactions is broken owing to the cleavage of surfaces, causing unsaturated O and Ti atoms to move outward (Supplementary Table 1). However, with the formation of Ti–F bonds, surface O and Ti atoms move inward and outward significantly, owing to the strong repulsive and attractive interactions

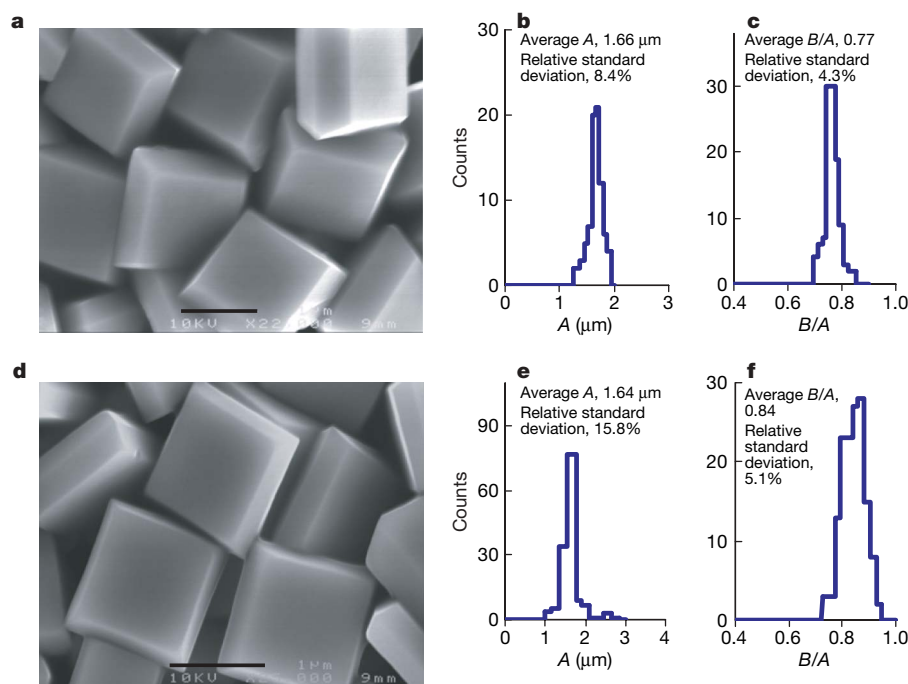


Figure 2 | SEM images and statistical data for the size and truncation degree of anatase single crystals. **a**, The morphology of anatase single crystals synthesized with 5.33 mM TiF_4 aqueous solution at 180°C for 14 h. **b**, The size (A) distribution of anatase single crystals in **a**. **c**, The degree of truncation (B/A) of anatase single crystals in **a**. **d–f**, As in **a–c**, but for

different synthesis conditions; the degree of truncation of anatase single crystals was controlled by halving the concentration of TiF_4 and extending the reaction time to 20 h. For both the size and the degree of truncation, the relative standard deviations of all statistical data were derived from counting more than 100 single crystals for each sample. Scale bars in **a** and **d**, $1 \mu\text{m}$.

of O–F and Ti–F bonds, respectively. This analysis is consistent with the calculated electronic structures (see Supplementary Fig. 3) because the Ti_{3d} and O_{2p} electrons can both interact strongly with the F_{2p} electron. A new balance can thereby be established between O–O/F–O repulsions and Ti–O/Ti–F attractions, which stabilizes Ti and O atoms on the surfaces.

As a further confirmation, no crystal facet control was observed in the absence of hydrofluoric acid, and only hollow spherical polycrystalline anatase particles were formed (see Supplementary Fig. 5)²⁸. Hydrofluoric acid is believed to have dual roles here: to retard hydrolysis of the titanium precursor and to reduce surface energy to promote the isotropic growth along the [010] and [100] axes^{18,20}, as illustrated in Fig. 1a–f. In comparison with the sample in Fig. 2a, the anatase single crystals obtained in shorter reaction times are of a smaller size and a higher degree of truncation (for example $A = 850$ nm and $B/A = 0.84$ at 8 h, depicted in Supplementary Fig. 6). According to the theoretical predictions (Fig. 1f), the B/A can be as great as 1.0 for the fully F-terminated surfaces; this important prediction indicates that ultrathin TiO_2 nanosheets may be synthesized.

Bright field images of transmission electron microscopy (TEM) and selected-area electron diffraction patterns confirm that each free-standing crystal shows single-crystal characteristics (Fig. 3a, b). The selected-area electron diffraction patterns can be indexed into diffraction spots of the [001] zone²⁸. The high-resolution TEM image (Fig. 3c) shows the (200) and (020) atomic planes with a lattice spacing of 1.89 Å and an interfacial angle of 90° (ref. 28). A corresponding fast-Fourier-transform-filtered TEM image of tetragonal atomic arrangement on the (001) surface is shown in Fig. 3d. The interfacial angle between two parallel faces and other surrounding faces is $68.3^\circ \pm 0.3^\circ$ on average (Supplementary Fig. 7); this value is identical to the theoretical value for the angle between the {001} and {101} facets of anatase⁵.

In the X-ray diffraction pattern shown in Fig. 4a, all the diffraction peaks match well with the crystal structure of the anatase TiO_2 phase

(space group $I4_1/amd$)²⁸. The X-ray photoelectron spectrum of F 1s core electrons (Fig. 4b) for the anatase single crystals clearly matches the model description in Fig. 1a–d; the measured binding energy is only 684.5 eV, which is a typical value for fluorinated TiO_2 systems such as TiOF_2 or the surface Ti–F species²⁹. Furthermore, the oxidation state of the Ti element in the same materials (Ti $2p_{3/2}$, binding energy 458.8 eV; Ti $2p_{1/2}$, binding energy 464.3 eV) is identical to that of bulk TiO_2 , as reported previously³⁰. From the X-ray photoelectron spectroscopy results, we can conclude that the atomic incorporation of F atoms, or their substitution for O atoms, in the anatase TiO_2 crystal lattice (doping) can be ruled out. These results strongly support the initial theoretical predictions and our explanation in terms of surface atoms; that is, the high F–Ti bonding energy significantly lowers the energy of the (001) surfaces, making them more stable than (101) surfaces in our reaction media. Importantly, the fluorinated surfaces of anatase single crystals can easily be made clean, fluorine-free surfaces by heating at 600°C for 90 min, without changing their crystal structure and morphology (Supplementary Fig. 8).

The well-defined, high-purity anatase single crystals synthesized in this work would be very useful as model single crystals for fundamental studies in surface science. Furthermore, high-purity anatase single crystals with a high percentage of reactive {001} facets have promising applications in solar cells, photonic and optoelectronic devices, sensors and photocatalysis. Our results illustrate the power of combining first-principle calculations and experimental techniques to achieve engineering of surface and crystallographic characteristics of crystalline materials.

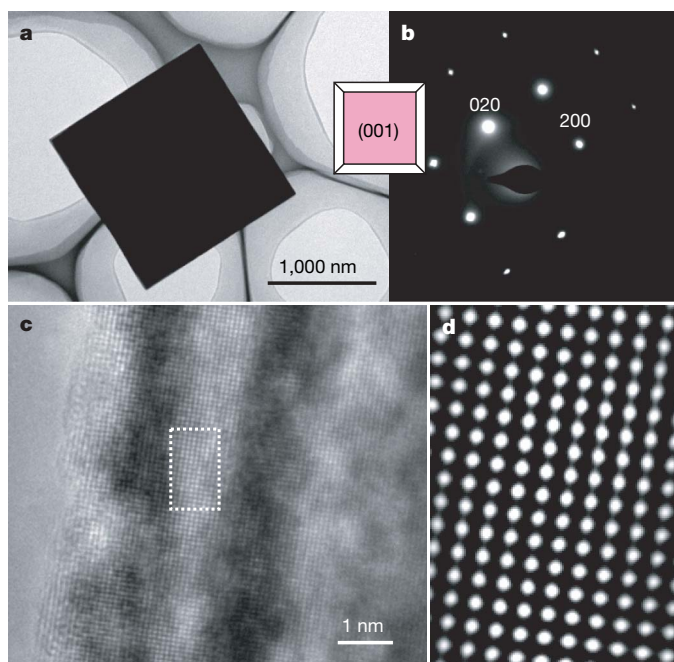


Figure 3 | Crystalline phase determination. **a**, Bright-field TEM image of a representative anatase single crystal recorded along the [001] axis. **b**, Presence of the anatase single-crystal phase can be confirmed by the square-symmetric selected-area electron diffraction pattern. Inset is a [001]-projected geometrical model of the anatase single crystals. **c**, High-resolution TEM image recorded from another anatase single crystal with [001] orientation. **d**, Fast-Fourier-transform-filtered TEM image recorded from the dotted rectangular area in **c**.

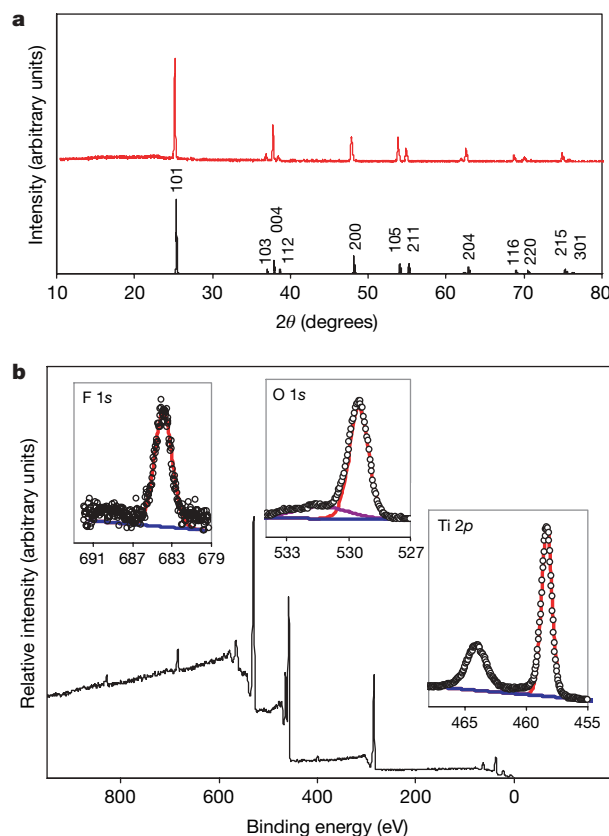


Figure 4 | Confirmation of the spatial distribution of F atoms in anatase single crystals. **a**, A representative X-ray diffraction pattern of the anatase single crystals we synthesized in this work (red), which is in good agreement with the calculated diffraction pattern of bulk anatase (black). θ , diffraction angle. **b**, X-ray photoelectron spectra of anatase single crystals, showing the four characteristic peaks of Ti, O, F and C. The insets have the same axes as the main panel. The open circles are the raw data of the X-ray photoelectron spectra, and the blue and red/purple lines represent the base line and fitted lines, respectively.

METHODS SUMMARY

Theoretical. Surface free energies were calculated using density functional theory within the generalized-gradient approximation. In each case, stoichiometric slab models (1×1) were employed, with all atoms being relaxed without any constraint.

Experimental. Titanium tetrafluoride aqueous solution (varying in concentration between 2.67 and 5.33 mM) and hydrofluoric acid (10 wt%) were used as the precursor and the crystallographic controlling agent, respectively, to prepare anatase TiO₂ single crystals. The reaction was carried out in a Teflon-lined autoclave at 180 °C for 2 to 20 h.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 26 November 2007; accepted 1 April 2008.

1. Tian, N., Zhou, Z. Y., Sun, S. G., Ding, Y. & Wang, Z. L. Synthesis of tetrahedral platinum nanocrystals with high-index facets and high electro-oxidation activity. *Science* **316**, 732–735 (2007).
2. Bikondoa, O. *et al.* Direct visualization of defect-mediated dissociation of water on TiO₂ (110). *Nature Mater.* **5**, 189–192 (2006).
3. Dulub, O. *et al.* Electron-induced oxygen desorption from the TiO₂ (011)- 2×1 surface leads to self-organized vacancies. *Science* **317**, 1052–1056 (2007).
4. Gong, X. Q., Selloni, A., Batzill, M. & Diebold, U. Steps on anatase TiO₂ (101). *Nature Mater.* **5**, 665–670 (2006).
5. Diebold, U. The surface science of titanium dioxide. *Surf. Sci. Rep.* **48**, 53–229 (2003).
6. Thomas, A. G. *et al.* Resonant photoemission of anatase TiO₂ (101) and (001) single crystals. *Phys. Rev. B* **67**, 035110 (2003).
7. Kavan, L., Grätzel, M., Gilbert, S. E., Klemenz, C. & Scheel, H. J. Electrochemical and photoelectrochemical investigation of single-crystal anatase. *J. Am. Chem. Soc.* **118**, 6716–6723 (1996).
8. Gong, X. Q. & Selloni, A. Reactivity of anatase TiO₂ nanoparticles: the role of the minority (001) surface. *J. Phys. Chem. B* **109**, 19560–19562 (2005).
9. Herman, G. S., Sievers, M. R. & Gao, Y. Structure determination of the two-domain (1×4) anatase TiO₂(001) surface. *Phys. Rev. Lett.* **84**, 3354–3357 (2000).
10. Lazzeri, M., Vittadini, A. & Selloni, A. Structure and energetics of stoichiometric TiO₂ anatase surfaces. *Phys. Rev. B* **63**, 155409 (2001).
11. Vittadini, A., Selloni, A., Rotzinger, F. P. & Grätzel, M. Structure and energetics of water adsorbed at TiO₂ anatase (101) and (001) surfaces. *Phys. Rev. Lett.* **81**, 2954–2957 (1998).
12. Vittadini, A., Casarin, M. & Selloni, A. Chemistry of and on TiO₂-anatase surfaces by DFT calculations: a partial review. *Theor. Chem. Acc.* **117**, 663–671 (2007).
13. Lazzeri, M. & Selloni, A. Stress-driven reconstruction of an oxide surface: the anatase TiO₂(001)-(1×4) surface. *Phys. Rev. Lett.* **87**, 266105 (2001).
14. Fujishima, A. & Honda, K. Electrochemical photolysis of water at a semiconductor electrode. *Nature* **238**, 37–38 (1972).
15. O'Regan, B. & Grätzel, M. A low-cost, high-efficiency solar cell based on dye-sensitized colloidal TiO₂ films. *Nature* **353**, 737–740 (1991).
16. Grätzel, M. Photoelectrochemical cells. *Nature* **414**, 338–344 (2001).
17. Barbé, C. J. *et al.* Nanocrystalline titanium oxide electrodes for photovoltaic applications. *J. Am. Ceram. Soc.* **80**, 3157–3171 (1997).
18. Penn, R. L. & Banfield, J. F. Morphology development and crystal growth in nanocrystalline aggregates under hydrothermal conditions: Insights from titania. *Geochim. Cosmochim. Acta* **63**, 1549–1557 (1999).
19. Zaban, A., Aruna, S. T., Tirosh, S., Gregg, B. A. & Mastai, Y. The effect of the preparation condition of TiO₂ colloids on their surface structures. *J. Phys. Chem. B* **104**, 4130–4133 (2000).
20. Jun, Y. W. *et al.* Surfactant-assisted elimination of a high energy facet as a means of controlling the shapes of TiO₂ nanocrystals. *J. Am. Chem. Soc.* **125**, 15981–15985 (2003).
21. Barnard, A. S. & Curtiss, L. A. Prediction of TiO₂ nanoparticle phase and shape transitions controlled by surface chemistry. *Nano Lett.* **5**, 1261–1266 (2005).
22. Chen, X. & Mao, S. S. Titanium dioxide nanomaterials: synthesis, properties, modifications, and applications. *Chem. Rev.* **107**, 2891–2959 (2007).
23. Izumi, F. The polymorphic crystallization of titanium (IV) oxide under hydrothermal conditions. II. The roles of inorganic anions in the nucleation of rutile and anatase from acid solutions. *Bull. Chem. Soc. Jpn* **51**, 1771–1776 (1978).
24. Berger, H., Tang, H. & Lévy, F. Growth and Raman spectroscopic characterization of TiO₂ anatase single crystals. *J. Cryst. Growth* **130**, 108–112 (1993).
25. Zmbov, K. F. & Margrave, J. L. Mass spectrometric studies at high temperatures. XVI. Sublimation pressures for TiF₃ (g) and the stabilities of TiF₂ (g) and TiF (g). *J. Phys. Chem.* **71**, 2893–2895 (1967).
26. Huber, K. P. & Herzberg, G. in *Molecular Spectra and Molecular Structure. IV. Constants of Diatomic Molecules* 642 (Van Nostrand Reinhold, New York, 1979).
27. Barnard, A. S. & Zapol, P. A model for the phase stability of arbitrary nanoparticles as a function of size and shape. *J. Chem. Phys.* **121**, 4276–4283 (2004).
28. Yang, H. G. & Zeng, H. C. Preparation of hollow anatase TiO₂ nanospheres via Ostwald ripening. *J. Phys. Chem. B* **108**, 3492–3495 (2004).
29. Yu, J. C., Yu, J., Ho, W., Jiang, Z. & Zhang, L. Effects of F[−] doping on the photocatalytic activity and microstructures of nanocrystalline TiO₂ powders. *Chem. Mater.* **14**, 3808–3816 (2002).
30. Lou, X. W. & Zeng, H. C. Complex α -MoO₃ nanostructures with external bonding capacity for self-assembly. *J. Am. Chem. Soc.* **125**, 2697–2704 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the Australian Research Council. H.G.Y. wishes to express his gratitude to the National University of Singapore, where the preliminary experimental work was carried out. The authors acknowledge Qiu Hong Hu for her help with statistical analysis.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to G.Q.L. (maxlu@uq.edu.au) or S.Z.Q. (s.qiao@uq.edu.au).

METHODS

Theoretical calculations. In each calculation, stoichiometric slab models (1×1) were used, consisting of nine atomic layers and a total of nine atoms for clean (001) surfaces and eight atomic layers and a total of 12 atoms for clean (101) surfaces. Each of these (001) and (101) surfaces contains one fivefold Ti atom terminated by an X atom (where X represents H, B, C, N, O, F, Si, P, S, Cl, Br or I). All atoms were relaxed without any constraint. All calculations were carried out using density functional theory within the generalized-gradient approximation³¹, with the exchange-correlation functional of ref. 32 (see also ref. 33). This was implemented in the Vienna *ab initio* simulation package^{34,35}, which spans reciprocal space with a plane-wave basis, in this case up to a kinetic energy cutoff of 450 eV. We used an $11 \times 11 \times 11$ Monkhorst–Pack *k*-point mesh for bulk anatase, an $11 \times 11 \times 1$ mesh for slabs, and a $3 \times 3 \times 3$ mesh for dimers of X for final calculations of energies. During the relaxations, all structures were relaxed to an energy convergence of 10^{-4} eV (equating to a force convergence of 10^{-2} eV Å⁻¹). In the case of slabs, the vacuum space is larger than 15 Å, and for the dimers of X a cubic unit cell with $a = b = c = 16$ Å was used. For isolated atoms (to correct the cohesive energy), a unit cell with $a = 15$ Å, $b = 16$ Å and $c = 17$ Å was used.

Synthesis of anatase TiO₂ single crystals. Hydrochloric acid (1.5 M) was used to adjust the pH of deionized water (1.01) to around 2.1. Titanium tetrafluoride (TiF₄, Aldrich Chemical) was then dissolved in this solution to a concentration of 0.04 M, which changed the pH to 1.8 (refs 36, 37). Finally, deionized water was used to adjust the concentration of TiF₄ aqueous solution to a concentration of 2.67 to 5.33 mM. In a typical synthesis, 30 ml of TiF₄ aqueous solution and 0.4 ml of hydrofluoric acid (10 wt%) were added to a Teflon-lined autoclave and a transparent mixture formed; the mixture was kept at 180 °C for 2 to 20 h in an oven. After completion of the reaction, the clear solution at the upper section was carefully removed by plastic dropper and a precipitate (TiO₂ product) formed at the bottom of the Teflon reactor. The solid products obtained were washed three times with deionized water (15 ml each time) and then dried at 100 °C or redispersed in deionized water for further characterization.

Fluorine removal from the surface of anatase TiO₂ single crystals. Typically, the powder samples of as-prepared anatase TiO₂ single crystals were heat treated in static air in a Muffle furnace at temperatures of 200 to 600 °C for 90 min with a ramping rate of 5 °C min⁻¹. The samples were then cooled to room temperature in the Muffle furnace, for further characterization.

Materials characterization. Crystallographic information of anatase TiO₂ single crystals was obtained with X-ray diffraction (Shimadzu XRD-6000, Cu K α radiation). Chemical compositions of anatase TiO₂ single crystals were analysed using X-ray photoelectron spectroscopy (Kratos Axis ULTRA incorporating a 165 mm hemispherical electron-energy analyser). All binding energies were referenced to the C 1s peak (285.0 eV) arising from adventitious carbon. Prior peak deconvolution, X-ray satellites and inelastic background (Shirley-type) were subtracted for all spectra. Morphology and crystal structure of anatase TiO₂ single crystals were examined using SEM (JEOL JSM6400F & JSM890) and TEM (Philips Tecnai T12 and T30F FEG Cryo AEM). Samples of anatase TiO₂ single crystals were dispersed in deionized water and dropped on a conductive SEM sample holder, or a carbon-coated copper grid with irregular holes for electron microscopy (SEM/TEM) analysis. X-ray photoelectron spectroscopy and X-ray diffraction samples were prepared by drying the sedimented particles overnight at 100 °C.

31. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev. B* **140**, A1133–A1138 (1965).
32. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
33. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).
34. Kresse, G. & Furthmüller, J. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
35. Kresse, G. & Furthmüller, J. Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
36. Yang, H. G. & Zeng, H. C. Creation of intestine-like interior space for metal-oxide nanostructures with a quasi-reverse emulsion. *Angew. Chem. Int. Ed.* **43**, 5206–5209 (2004).
37. Yang, H. G. & Zeng, H. C. Synthetic architectures of TiO₂/H₂Ti₅O₁₁·H₂O, ZnO/H₂Ti₅O₁₁·H₂O, ZnO/TiO₂/H₂Ti₅O₁₁·H₂O and ZnO/TiO₂ nanocomposites. *J. Am. Chem. Soc.* **127**, 270–278 (2005).

LETTERS

Snowball Earth termination by destabilization of equatorial permafrost methane clathrate

Martin Kennedy¹, David Mrofka¹ & Chris von der Borch²

The start of the Ediacaran period is defined by one of the most severe climate change events recorded in Earth history—the recovery from the Marinoan ‘snowball’ ice age, ~635 Myr ago (ref. 1). Marinoan glacial-marine deposits occur at equatorial palaeolatitudes², and are sharply overlain by a thin interval of carbonate that preserves marine carbon and sulphur isotopic excursions of about –5 and +15 parts per thousand, respectively^{3–5}; these deposits are thought to record widespread oceanic carbonate precipitation during postglacial sea level rise^{1,3,4}. This abrupt transition records a climate system in profound disequilibrium^{3,6} and contrasts sharply with the cyclical stratigraphic signal imparted by the balanced feedbacks modulating Phanerozoic deglaciation. Hypotheses accounting for the abruptness of deglaciation include ice albedo feedback³, deep-ocean out-gassing during post-glacial oceanic overturn⁷ or methane hydrate destabilization^{8–10}. Here we report the broadest range of oxygen isotope values yet measured in marine sediments (–25‰ to +12‰) in methane seeps in Marinoan deglacial sediments underlying the cap carbonate. This range of values is likely to be the result of mixing between ice-sheet-derived meteoric waters and clathrate-derived fluids during the flushing and destabilization of a clathrate field by glacial meltwater. The equatorial palaeolatitude implies a highly volatile shelf permafrost pool that is an order of magnitude larger than that of the present day. A pool of this size could have provided a massive biogeochemical feedback capable of triggering deglaciation and accounting for the global postglacial marine carbon and sulphur isotopic excursions, abrupt unidirectional warming, cap carbonate deposition, and a marine oxygen crisis. Our findings suggest that methane released from low-latitude permafrost clathrates therefore acted as a trigger and/or strong positive feedback for deglaciation and warming. Methane hydrate destabilization is increasingly suspected as an important positive feedback to climate change^{11–13} that coincides with critical boundaries in the geological record^{14,15} and may represent one particularly important mechanism active during conditions of strong climate forcing.

Authigenic carbonates within the Reynella Member of the glacial Elatina Formation occur in coastal cliff exposures of the Marinoan type section in South Australia (Fig. 1 a–d) and provide constraints on the magnitude and timing of methane influence during deglaciation. The Reynella Member is part of a transgressive littoral succession that deepens into the shelfal deposits of the Seacliff sandstone and interbedded cap carbonate of the Nuccaleena Formation¹⁶ (Fig. 1f). The Nuccaleena Formation is the Marinoan cap carbonate that houses the global stratotype section and point for the Ediacaran period, and is widely believed to provide a global tie point shared by the integrated sea level rise driven by global deglaciation⁴. On the basis of the low palaeolatitude², minimal glacial evidence and transgressive sediments, we interpret deposition of the Reynella siltstone

to coincide with local ice withdrawal preceding global eustatic rise recorded by the Nuccaleena Formation (Fig. 1f).

These carbonates form an irregular plumbing system containing sub-vertical chimneys 1–2 m in diameter and branching pipes (Fig. 1a) that coalesce into hard grounds. Microcrystalline dolomite and calcite cement fills pores in sandstone and forms up to 60% of the sediment, indicating precipitation before compaction. In plan view, spar-filled tubes are concentrated within the cores of 1-m-diameter chimneys ringed by cemented sandstone. Morphologies include microcrystalline pore-filling calcite and Fe-rich dolomite with pyrite crystals >10 µm indicative of sulphate reduction, acicular cements (formerly aragonite) fringing syndimentary breccia clasts of carbonate-cemented sandstone, spar-lined vugs and tubes with Mn-cement coatings, sparry calcite-filled veins often forming a honeycomb crust morphology, and isolated centimetre-scale bladed crystal fans of calcite pseudomorphs after barite. Calcite spar is present in early veins, thin vertical tubes and vugs, and within hard grounds. Both dolomite and calcite cements and calcite spar comprise clasts eroded and reworked by currents within channelized, cross-bedded sand and breccia deposits (Fig. 1b). These indicate a dynamic system, including erosion with re-cementation in shallow seafloor sediments rather than a more typical history of cementation during burial. Breccias of similar composition occur as metre-scale, massive, vertically oriented cross-cutting fluidized beds (Fig. 1c, d) that erupted on the sea floor and were reworked by seafloor currents into cross-bedded channel-fill. Sedimentary features indicative of vertical fluid or gas migration, overpressure, seafloor reworking and complex seafloor paragenesis are diagnostic of a methane cold-seep origin^{17,18}. This complex morphology contrasts with the regionally extensive laminated dolomicrite of the overlying Nuccaleena Formation and the sparry calcite that fills cross-cutting fractures aligned with regional folding.

Authigenic cements within these morphologies show the broadest range of $\delta^{18}\text{O}$ values ever reported from marine sediments¹⁹, ranging from +12‰ to –25‰ (PDB) at centimetre-scale resolution (Fig. 2; Supplementary Table 1). Potential controls on oxygen isotope values include precipitating fluid temperature, fractionation associated with crystal formation, and the composition of the source pore fluid. The heaviest $\delta^{18}\text{O}$ values (Fig. 2; +12‰) occur in pore-filling dolomite cement and are too enriched in ^{18}O to be attributed to conventional mechanisms such as direct dolomite precipitation from sea water at 0°C (ref. 17). Values over 10‰ in marine sediments are known only from methane cold-seep environments where ^{18}O is enriched in pore fluids during methane clathrate formation and dolomite precipitation results from sulphate oxidation of methane^{17,20}. Values of $\delta^{13}\text{C}$ that range from –10‰ to +10‰ PDB (Fig. 2; Supplementary Table 1) within Reynella cements are also common within cold-seep-associated cements and attributed to

¹Department of Earth Science, University of California, Riverside, California 92521, USA. ²School of Chemistry, Physics and Earth Sciences, Flinders University, GPO Box 2100 Adelaide, South Australia, 5001 Australia.

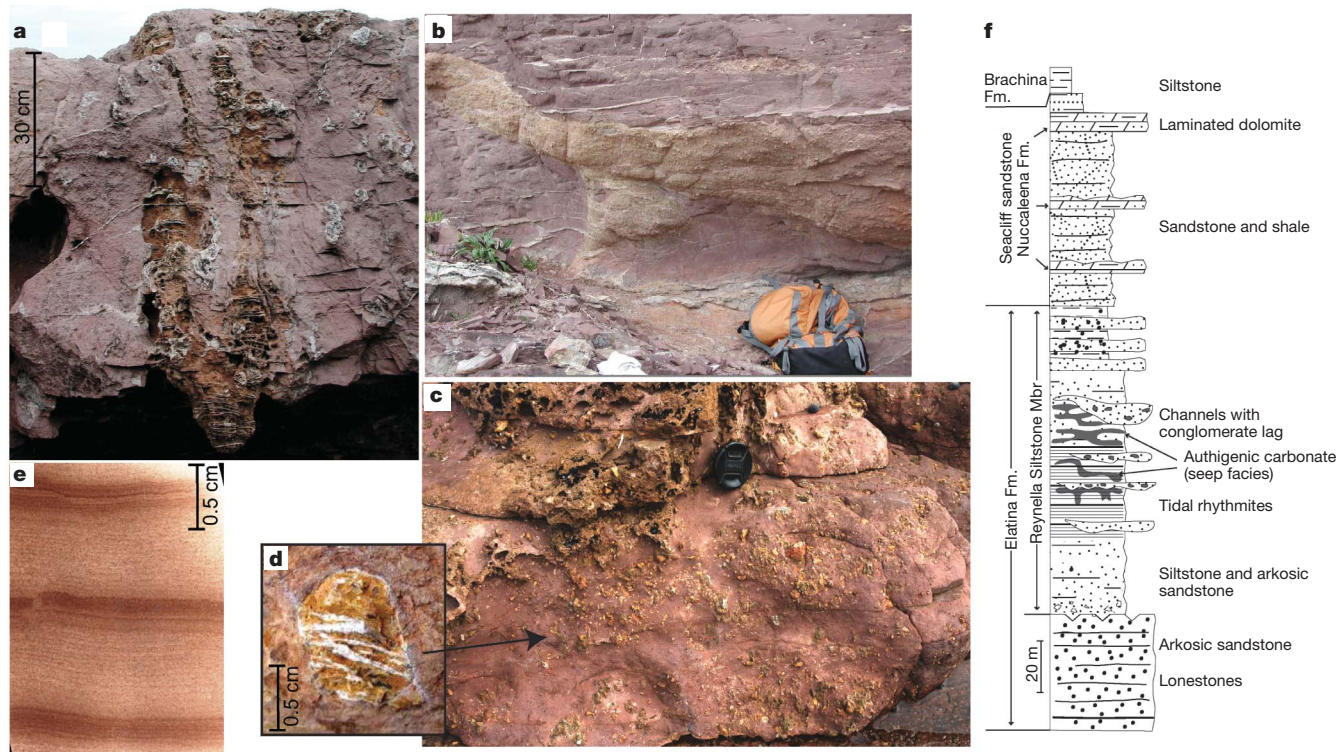


Figure 1 | Photographic examples and stratigraphic context for Reynella seep carbonates. Distinctive sedimentary facies include vertical cemented chimneys rising through coarse sandstone with sparry cavity cements (a); shallow channels (backpack for scale) with basal lags of authigenic carbonate clasts eroded from exhumed seeps (b); and breccia beds comprising cemented sandstone (5-cm lens cap for scale) resulting from

fluidization (c, d). A thin section of tidal rhythmites (e) shows well-developed neap–spring bundles of couplets from samples collected between authigenic carbonate hard grounds. Schematic section (f) shows the association of tidal, seep and stratigraphic units. Section and photos located between Hallet Cove and Marino Rocks around 36° 03' 45" S, 138° 30' 05" E.

concurrent methanotrophy and mixing between isotopically heavy methanogenic CO_2 and marine bicarbonate^{17,18,20}.

Values of $\delta^{18}\text{O}$ for calcite cements within the Reynella Member define a mixing line with a highly depleted meteoric endmember (-25‰ PDB; $R^2 = 0.84$). Although low $\delta^{18}\text{O}$ values can occur during

metamorphism by equilibration at $>180^\circ\text{C}$, a metamorphic origin is incompatible with the presence of temperature-sensitive detrital, dioctahedral-smectite in the laminated mud fraction of the host rock which would have converted to illite above 100°C (Supplementary Fig. 1; Supplementary Table 2). Thermal equilibration is inconsistent with (1) the microcrystalline morphology of the cement, (2) the isotopic heterogeneity on millimetre scales, and (3) the relatively heavy $\delta^{18}\text{O}$ values of the Nuccaleena cap carbonate and late-stage (tectonic) cross-cutting spar fill fractures (Fig. 2). The isotopic signature of the most depleted cements (-25‰) requires porewater values as low as about -28 to -30‰ (see Supplementary Information), which, in the modern hydrological cycle, is unique to meltwater released from glacial ice²¹. In the geological record, these low values are found only in thermally altered carbonate spar¹⁹ or formed from glacial meltwater in ancient carbonates^{22,23}.

The mixing relation connecting enriched and depleted $\delta^{18}\text{O}$ endmembers (Fig. 2), sub-centimetre spatial association and syndeformational timing relations between stable isotopic endmembers indicates formation of seeps in a shallow environment in which meteoric fluid derived from an ice sheet mixed with ^{18}O -enriched clathrate-derived fluids. Such an environment occurs in shelf settings where the hydraulic drive of a melting ice sheet charges the water table, flushing and destabilizing clathrate fields.

Sedimentary evidence indicates that the Reynella seeps formed syndeformationally in high-energy, intertidal, sandstones. The authigenic carbonate occurs within a succession containing superbly preserved tidal bundles (Fig. 1e) including well-developed spring and neap cycles²⁴, broad (2 m), shallow (1 m) channels (Fig. 1b), reactivation surfaces, bimodal palaeocurrents within cross-bedded sands, and herringbone and flaser lamination, all indicative of a high-energy, macrotidal environment^{16,24} unique to coastal or estuarine settings. Reworking of authigenic carbonate in channels followed by

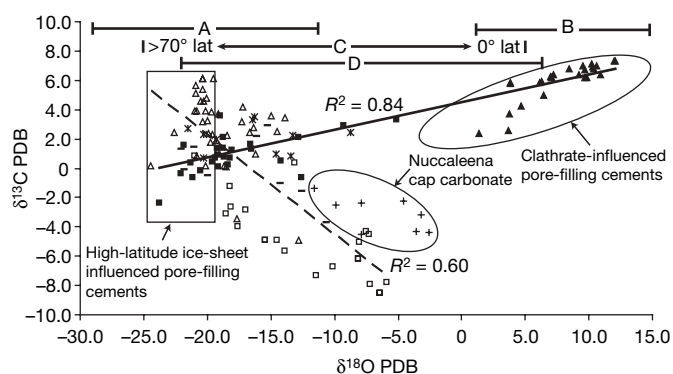


Figure 2 | Values of $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ for Nuccaleena Formation and Reynella Member cements and veins. Open triangles, calcite-cemented siltstone; filled triangles, dolomite-cemented siltstones; open squares, spar-filled cavities; filled squares, early calcite spar; asterisks, early spar-filled tubes; dashes, late cross-cutting veins; crosses, Nuccaleena cap carbonate. Upper trend comprises filled symbols and asterisks, lower trendline includes open symbols. Later phases (crosses and dashes) not included in trend lines. Horizontal bars represent ranges of $\delta^{18}\text{O}$ values from (A) isotopically light meltwater-related carbonates^{23,24}; (B) clathrate enrichment of ^{18}O in seep carbonates^{18,19,21}; (C) predicted $\delta^{18}\text{O}$ values from carbonate precipitated in 0°C meteoric water from $>70^\circ$ latitude (-26‰ PDB calcite from -30‰ SMOW water) to low latitude ($+4\text{‰}$ PDB from -2‰ SMOW water)²²; and (D) reported range of $\delta^{18}\text{O}$ values for all marine carbonate rocks³⁰.

subsequent cementation (Fig. 1b) indicates that seep hard grounds formed throughout the tidally influenced deposition of sand.

Methane clathrates form when methane produced by thermogenic or biogenic breakdown of organic matter at depth migrates upwards to stabilize more than 200 m below the sea floor or land surface in a narrow range of pressure and temperature^{12,25}. To be stable, marine clathrates require both sea water at 0 °C and the additional pressure derived from more than 250 m of water¹². Because the Reynella Member consists of intertidal deposits within a deepening succession midway between the preceding glacial sealevel lowstand and the shelfal deposits of the Seaciff sandstone/Nuccaleena dolomite, the water could not have been deep enough to stabilize clathrates. The stabilization thus must have resulted either from the pressure of a shelf-borne ice sheet or from the permafrost conditions (<0 °C) indicated by local evidence of frost wedging²⁶. Permafrost clathrates that accumulated beneath the broad exposed shelves of the Arctic during the Pleistocene are now destabilizing as a result of post-glacial flooding and global warming^{11,12,25}. Similarly, retreat of the Marinoan ice sheet would have destabilized methane clathrates by (1) reducing hydrostatic pressure, (2) warming permafrost shelves with flooding sea water, and (3) flushing clathrate fields with meteoric water derived from meltwater discharge^{9,22,23}.

The extension of a volatile permafrost clathrate pool to equatorial palaeolatitudes has important implications for the role of methane in the climate system. Initial warming or, more simply, inherent ice-sheet instability²⁷, might have been enough to trigger local methane destabilization through meteoric flushing of the permafrost clathrate field on the Marinoan shelf, particularly in this low-latitude setting² where it was most sensitive to subtle external forcing. Once initiated, clathrate destabilization provides a positive feedback through greenhouse warming caused by methane and CO₂, and is amplified by further destabilization in progressively higher latitudes^{9,11} of permafrost and eventually marine clathrates. Extension of the Holocene permafrost reservoir, occurring over the Arctic shelves and estimated²⁵ at 400 gigatonnes, to total present-day shelf area (including low latitudes) could increase the potential size to ~3,000 Gt (see Supplementary Information). This estimate is conservative because it does not include the extensive intracratonic basinal areas characteristic of the Neoproterozoic (for example⁹), nor does it consider destabilization of the more conventional and tenfold larger marine clathrate reservoir²⁵. It is, however, comparable to past gas hydrate events with estimated releases of ~5,000 Gt in the Toarcian¹⁴ or ~1,000–2,100 Gt in the Palaeocene¹⁵ estimated from marine excursions of –7‰ and –3‰ $\delta^{13}\text{C}$, respectively. Permafrost clathrate destabilization could readily provide a trigger for feedbacks to deglaciation such as albedo, marine clathrate destabilization or carbonate reallocation^{3,28}.

Alkalinity increase capable of accounting for cap carbonate deposition could arise on shorter timescales (<10⁴ yr) from sulphate oxidation of methane following exhaustion of dissolved oxygen in pore and sea water¹⁰ and on longer timescales (>10⁵ yr) from oxidation of methane to CO₂ in the atmosphere driving increased continental chemical weathering. The duration of methane release would most likely be controlled by pulsed marine clathrate destabilization similar to that in the early Toarcian¹⁴, which occurred over ~200 kyr. Like other later methane destabilization events¹⁴, release of 3,000 Gt of methane (~2 × 10¹⁷ moles of carbon, excluding marine clathrates) into sea water over about 10³–10⁴ yr could substantially influence the redox state of parts of the ocean, significantly depleting the dissolved oxygen pool and consuming marine sulphate which maintained a probable stock of less than 4.5 × 10¹⁶ mole O₂ (ref. 29) and 5 × 10¹⁷ mole SO₄ (refs 2–5) (see Supplementary Information). Significant sulphate oxidation of methane within pore and sea water predicts the observed 15‰ positive shift in marine sulphate $\delta^{34}\text{S}$ values⁵. The release of 3,000 Gt of –60‰ $\delta^{13}\text{C}$ methane from the permafrost reservoir could alone account for a shift of more than –3‰ (see Supplementary Information) of the enigmatic negative

$\delta^{13}\text{C}$ excursion in marine carbonates that defines the base of the Ediacaran period⁴.

The distinctive features of Marinoan deglaciation that define the base of the Ediacaran period⁴ can be attributed to the effects of permafrost methane clathrate destabilization. In contrast to the balanced feedbacks and progressive glacial–interglacial cycles of Cenozoic deglaciation, the violent opening of the highly volatile shelf-permafrost methane clathrate pool could act as a trigger to catastrophic climate and biogeochemical reorganization of the Earth system, abruptly bringing the long-lived and icy Cryogenian period to a close and setting the stage for the appearance of metazoans and dominance of a new Earth system. This event both identifies the range of function of the climate system, and demonstrates a mechanism activated by strong climate forcing not unlike projected future effects of atmospheric CO₂.

METHODS SUMMARY

Detailed methods are given in the Supplementary Information. We characterized the petrography and mineralogy by X-ray diffraction, petrographic microscopy and back scattered electron microscopy. Isotopic results were obtained on a VG Isogas Prism II, with reproducibility of $\delta^{18}\text{O}$ -enriched samples confirmed on a Thermo Delta-V Plus and at the University of California, Davis, and agree within 0.1‰. $\delta^{18}\text{O} = [(^{18}\text{O}/^{16}\text{O})_{\text{sample}}/(^{18}\text{O}/^{16}\text{O})_{\text{PDB}}] - 1$; $\delta^{13}\text{C} = [(^{13}\text{C}/^{12}\text{C})_{\text{sample}}/(^{13}\text{C}/^{12}\text{C})_{\text{PDB}}] - 1$.

Received 24 September 2007; accepted 18 March 2008.

1. Fairchild, I. J. & Kennedy, M. J. Neoproterozoic glaciation in the Earth system. *J. Geol. Soc. Lond.* **164**, 895–921 (2007).
2. Sohl, L. E., Christie-Blick, N. & Kent, D. V. Paleomagnetic polarity reversals in Marinoan (ca. 600 Ma) glacial deposits of Australia; implications for the duration of low-latitude glaciation in Neoproterozoic time. *Geol. Soc. Am. Bull.* **111**, 1120–1139 (1999).
3. Hoffman, P. F., Kaufman, A. J., Halverson, G. P. & Schrag, D. P. A Neoproterozoic snowball Earth. *Science* **281**, 1342–1346 (1998).
4. Knoll, A. H., Walter, M. R., Narbonne, G. M. & Christie-Blick, N. The Ediacaran Period: a new addition to the geologic time scale. *Lethaia* **39**, 13–30 (2006).
5. Halverson, G. P. & Hurtgen, M. T. Ediacaran growth of the marine sulfate reservoir. *Earth Planet. Sci. Lett.* **263**, 32–44 (2007).
6. Allen, P. A. & Hoffman, P. F. Extreme winds and waves in the aftermath of a Neoproterozoic glaciation. *Nature* **433**, 123–127 (2005).
7. Grotzinger, J. P. & Knoll, A. H. Anomalous carbonate precipitates; is the Precambrian the key to the Permian? *Palaos* **10**, 578–596 (1995).
8. Jiang, G., Kennedy, M. J. & Christie-Blick, N. Stable isotopic evidence for methane seeps in Neoproterozoic postglacial cap carbonates. *Nature* **426**, 822–826 (2003).
9. Shields, G. A., Deynoux, M., Strauss, H., Paquet, H. & Nahon, D. Barite-bearing cap dolostones of the Taoudeni Basin, northwest Africa; sedimentary and isotopic evidence for methane seepage after a Neoproterozoic glaciation. *Precamb. Res.* **153**, 209–235 (2007).
10. Kennedy, M. J., Christie-Blick, N. & Sohl, L. E. Are Proterozoic cap carbonates and isotopic excursions a record of gas hydrate destabilization following Earth's coldest intervals? *Geology* **29**, 443–446 (2001).
11. Nisbet, E. G. The end of the ice age. *Can. J. Earth Sci.* **27**, 148–157 (1990).
12. MacDonald, G. J. Role of methane clathrates in past and future climates. *Clim. Change* **16**, 247–281 (1990).
13. Archer, D. Methane hydrate stability and anthropogenic climate change. *Biogeosciences* **4**, 521–544 (2007).
14. Kemp, D. B., Coe, A. L., Cohen, A. S. & Schwark, L. Astronomical pacing of methane release in the Early Jurassic period. *Nature* **437**, 396–399 (2005).
15. Dickens, G. R., O'Neil, J. R., Rea, D. K. & Owen, R. M. Dissociation of oceanic methane hydrate as a cause of the carbon isotope excursion at the end of the Paleocene. *Paleoceanography* **10**, 965–971 (1995).
16. Preiss, W. V. The Adelaide Geosyncline, late Proterozoic stratigraphy, sedimentation, palaeontology, and tectonics. *Bull. Geol. Surv. S. Aust.* **53**, 1–438 (1987).
17. Greinert, J., Bohrmann, G. & Suess, E. Gas hydrate-associated carbonates and methane-venting at Hydrate Ridge; classification, distribution, and origin of authigenic lithologies. *Geophys. Monogr.* **124**, 99–113 (2000).
18. Aiello, I. W., Garrison, R. E., Moore, J. C., Kastner, M. & Stakes, D. S. Anatomy and origin of carbonate structures in a Miocene cold-seep field. *Geology* **29**, 1111–1114 (2001).
19. Shields, G. & Veizer, J. Precambrian marine carbonate isotope database versions 1.1. *Geochem. Geophys. Geosyst.* **3**, 1–12 (2002).
20. Ritger, S., Carson, B. & Suess, E. Methane-derived authigenic carbonates formed by subduction-induced pore-water expulsion along the Oregon/

- Washington margin; with Suppl. Data 87–02. *Geol. Soc. Am. Bull.* **98**, 147–156 (1987).
21. Bowen, G. J. & Wilkinson, B. Spatial distribution of $\delta^{18}\text{O}$ in meteoric precipitation. *Geology* **30**, 315–318 (2002).
 22. Herbert, C. T. & Compton, J. S. Depositional environments of the Lower Permian Dwyka Diamictite and Prince Albert Shale inferred from the geochemistry of early diagenetic concretions, southwest Karoo Basin, South Africa. *Sedim. Geol.* **194**, 263–277 (2007).
 23. Himmler, T., Freiwald, A., Stollhofen, H. & Peckmann, J. Late Carboniferous hydrocarbon-seep carbonates from the glaciomarine Dwyka Group, southern Namibia. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **257**, 185–197 (2008).
 24. Williams, G. E. in *Clastic Tidal Sedimentology* (eds Smith, D. G., Reinson, G. E., Zaitlin, B. A. & Rahmani, R. A.) 161–178 (Canadian Society of Petroleum Geologists, Calgary, 1991).
 25. Kvenvolden, K. A. Methane hydrate; a major reservoir of carbon in the shallow geosphere? *Chem. Geol.* **71**, 41–51 (1988).
 26. Williams, G. E. Precambrian permafrost horizons as indicators of palaeoclimate. *Precamb. Res.* **32**, 233–242 (1986).
 27. MacAyeal, D. R. Binge/Purge oscillations of the Laurentide ice sheet as a cause of the North Atlantic Heinrich events. *Paleoceanography* **8**, 775–784 (1993).
 28. Ridgwell, A. J., Kennedy, M. J. & Caldeira, K. Carbonate deposition, climate stability, and Neoproterozoic ice ages. *Science* **302**, 859–862 (2003).
 29. Canfield, D. E., Poulton, S. W. & Narbonne, G. M. Late-Neoproterozoic deep-ocean oxygenation and the rise of animal life. *Science* **315**, 92–95 (2007).
 30. Hoefs, J. *Stable Isotope Geochemistry* 55 (Springer, Berlin, 2004).
- Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.
- Acknowledgements** We thank the NSF and NASA Exobiology for funding, A. Derkowski and T. Bristow for assistance with clay mineralogy, D. Winter for replicate isotopic analyses and I. Fairchild for a review.
- Author Contributions** M.J.K. collected samples and wrote the manuscript; D.D.M. collected samples, performed geochemical analyses and wrote the supplemental section; C.v.d.B. collected samples. All authors discussed the results and commented on the manuscript.
- Author Information** Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.K. (martink@mail.ucr.edu).

LETTERS

A large discontinuity in the mid-twentieth century in observed global-mean surface temperature

David W. J. Thompson¹, John J. Kennedy², John M. Wallace³ & Phil D. Jones⁴

Data sets used to monitor the Earth's climate indicate that the surface of the Earth warmed from ~1910 to 1940, cooled slightly from ~1940 to 1970, and then warmed markedly from ~1970 onward¹. The weak cooling apparent in the middle part of the century has been interpreted in the context of a variety of physical factors, such as atmosphere–ocean interactions and anthropogenic emissions of sulphate aerosols². Here we call attention to a previously overlooked discontinuity in the record at 1945, which is a prominent feature of the cooling trend in the mid-twentieth century. The discontinuity is evident in published versions of the global-mean temperature time series¹, but stands out more clearly after the data are filtered for the effects of internal climate variability. We argue that the abrupt temperature drop of ~0.3 °C in 1945 is the apparent result of uncorrected instrumental biases in the sea surface temperature record. Corrections for the discontinuity are expected to alter the character of mid-twentieth century temperature variability but not estimates of the century-long trend in global-mean temperatures.

The time series of global-mean surface temperatures (T_G) reflects the influence of both anthropogenically induced and internally driven climate variability. A common method used to identify the signature of anthropogenic emissions in T_G is to compare the observed temperature variability with the output of climate change simulations. An alternative approach used to isolate the anthropogenic signal in T_G is to filter out the variance associated with internally driven climate phenomena and then interpret the residual time series. Here we exploit the latter technique to filter out the variance in T_G associated with two prominent internally driven climate phenomena: (1) the El Niño/Southern Oscillation (ENSO); and (2) temperature advection over the high latitudes of the Northern Hemisphere during winter.

The data and details of the filtering method are described in Methods. Briefly, the signal of ENSO in T_G is modelled as the damped thermodynamic response of the tropics to anomalous sea–air fluxes of sensible and latent heat in the dynamically active cold-tongue region of the eastern tropical Pacific³. The resulting ENSO index time series (T_{ENSO} ; Fig. 1) is analogous to a lagged and low-pass-filtered time series of sea surface temperatures (SSTs) averaged over the eastern tropical Pacific, and provides a substantially better rendition of the ENSO signal in globally averaged temperatures than the more commonly used cold-tongue index (Table 1).

The signal of temperature advection over the high latitudes of the Northern Hemisphere is manifested in the so-called ‘cold oceans–warm land’ (COWL) pattern⁴. During months with abnormally strong westerly winds at the surface, there is enhanced advection of relatively warm marine air masses over the colder continents and cold continental air masses over the warmer oceans. The continents have a lower heat capacity than the oceans, hence the warming of the land exceeds

the cooling of the ocean, and the global-mean temperature for that month is anomalously high. Months with abnormally weak surface westerlies are marked by global-mean temperature anomalies in the opposite sense. The calculation of the COWL pattern and its associated index time series is described in Methods. The COWL index time series (T_{COWL}) accounts for a substantial amount of the month-to-month weather-related ‘noise’ in T_G but also has weak secular variability due in part to trends in the atmospheric circulation (Fig. 1).

The influences of ENSO and the COWL pattern on surface temperatures were removed by subtracting the linearly fitted T_{ENSO} and T_{COWL} index time series from T_G (Methods). The resulting residual global-mean temperature time series (T_G^{residual}) is shown at the bottom of Figs 1 and 2. Filtering out ENSO and the COWL pattern reduces substantially the amount of interannual and month-to-month variance in T_G without reducing its temporal resolution. Consequently, the residual time series provides a cleaner rendition of the interdecadal variability in the time series of twentieth-century global-mean temperatures while retaining and increasing the prominence of numerous discrete drops in it. Most of the more prominent drops that are apparent in T_G^{residual} coincide with large tropical

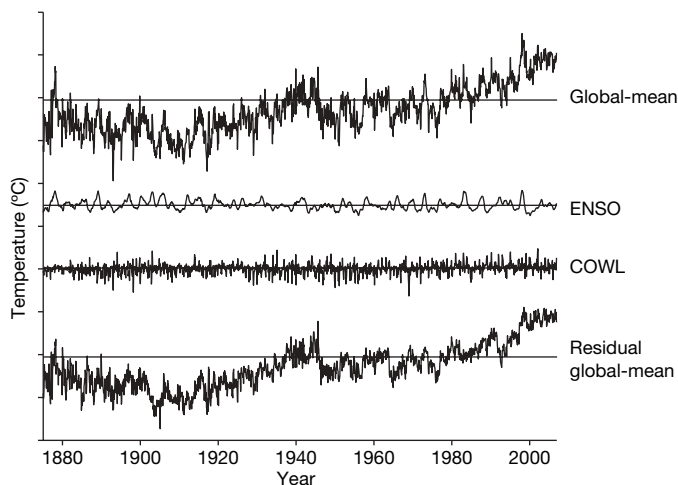


Figure 1 | Filtering out the ENSO and COWL signatures from the global-mean temperature time series. From top to bottom: the time series of global-mean, monthly-mean surface temperature anomalies based on the HadCRUT3 combined SST and land surface air temperature data (T_G ; data described in Methods); the contribution of T_{ENSO} to T_G ; the contribution of T_{COWL} to T_G ; the residual global-mean time series found by removing from T_G the linear contributions of T_{ENSO} and T_{COWL} . Vertical axis shows temperature anomalies; tickmarks indicate steps of 0.5 °C. Horizontal lines denote the mean of the respective time series for the 1961–90 period. The calculation of the ENSO and COWL index time series and the linear fitting technique are described in Methods.

¹Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado 80523, USA. ²Met Office Hadley Centre, Exeter EX1 3PB, UK. ³Department of Atmospheric Sciences, University of Washington, Seattle, Washington 98195, USA. ⁴Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

Table 1 | Correlations between indicated time series

| | T_G | T_G^{land} | T_G^{ocean} |
|-------------------|-------------|---------------------|----------------------|
| Cold-tongue index | 0.28 | 0.16 | 0.39 |
| T_{ENSO} | 0.38 | 0.30 | 0.45 |
| COWL | 0.41 | 0.50 | -0.02 |

The cold-tongue index corresponds to SST anomalies averaged over the region 5° N–5° S and 180–90° W. The analyses used to calculate the T_{ENSO} and COWL index time series and the data used to calculate the global-mean combined land+ocean (T_G), land (T_G^{land}) and ocean (T_G^{ocean}) time series are described in Methods. Correlations are based on detrended monthly-mean data 1950–2006. Bold font denotes correlations that are significant at the 95% level assuming one degree of freedom per year.

volcanic eruptions (Fig. 2; solid vertical lines). However, the most pronounced drop occurs in late 1945 and is not associated with any known climate phenomenon. The signal of the volcanic eruptions in the residual data is discussed elsewhere (D.W.J.T. *et al.*, manuscript in preparation); here we focus on the discontinuity in late 1945.

The discontinuity in global-mean surface temperatures in late 1945 is evident in the unfiltered global-mean time series, but its prominence and unique character are highlighted by the removal of the ENSO and COWL-related variability (Fig. 2). The amplitude of the step is substantial: temperatures dropped by $\sim 0.3^\circ\text{C}$ during the ~ 6 -month period following August 1945, thus the amplitude of the drop is roughly 40% as large as the $\sim 0.75^\circ\text{C}$ rise in T_G from 1900 to 2006, and is larger than the drop in T_G following the June 1991 eruption of Mt Pinatubo (Fig. 2 bottom time series). Clearly, the step has a profound impact on the historical record of twentieth-century surface temperatures.

The step in late 1945 does not appear to be related to any known physical phenomenon. No substantial volcanic eruptions were reported at the time, and the nuclear explosions over Hiroshima and Nagasaki are estimated to have had little effect on global-mean temperatures: ~ 100 Hiroshima-sized explosions are predicted to lead to a global-mean cooling of $\sim 1.25^\circ\text{C}$ (ref. 5), thus two such explosions might be expected to lead to a cooling of less than 0.03°C . Furthermore, ocean and land areas should both respond to an external forcing, but the step is only apparent in SSTs (Fig. 3). The global-mean land time series does not exhibit warming from the middle of the century until about 1980, but there is no large discrete drop in late 1945 in the unfiltered land series and only an indistinct drop in the residual land series (Fig. 3b). As is the case for the global-mean time series in Fig. 2, the drop is apparent in the unfiltered global-mean SST time series but is highlighted after filtering out the effects of internal climate variability.

Why did global-mean SSTs drop so rapidly in late 1945? At the time, SST data were sampled on board ships using a variety of measurement techniques, including measurements taken from insulated and uninsulated buckets and engine room intakes (see refs 6–9 and

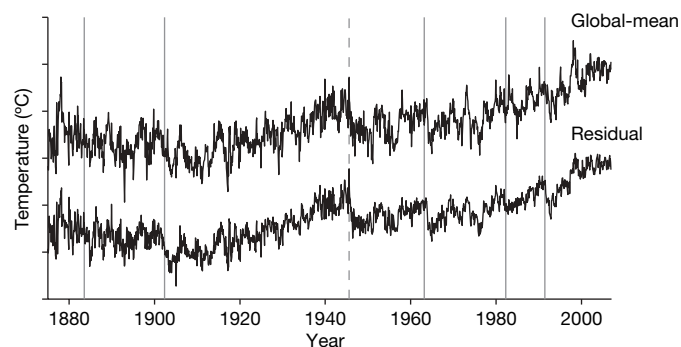


Figure 2 | The original (that is, unfiltered) and residual global-mean temperature time series duplicated from Fig. 1. The solid vertical lines denote volcano eruption dates: from left to right, Mts Krakatoa, Santa Maria, Agung, El Chichón and Pinatubo. The dashed vertical line denotes the month of August 1945. Vertical axis shows temperature anomalies; tickmarks indicate steps of 0.5°C .

references therein). Different measurement techniques are prone to different physical biases. For example, bucket measurements are affected by the exchange of sensible, latent and radiant heat with the surrounding environment, and are often biased cool relative to the actual temperature. Engine room intake measurements are influenced by the proximity of the ship's engine and the depth of the intake water, and are most often biased warm^{6,9}. The effect of changes in instrumentation on SST measurements can be corrected only with knowledge of both the physical biases and a time history of the way in which measurements were taken.

The results shown here are based on the current version of the UK Met Office Hadley Centre SST data set (HadSST2; ref. 9). Like all historical SST products, the HadSST2 data are derived from the database of raw (that is, uncorrected) marine observations provided by the International Comprehensive Ocean-Atmosphere Data Set (ICOADS; ref. 10), or one of its predecessors. The SST data in the ICOADS archive are affected by numerous changes in instrumentation. The HadSST2 data have been corrected for the widespread use of uninsulated buckets in the nineteenth and early twentieth century, and for the large change in instrumentation which occurred when US Merchant Marine ships switched from uninsulated bucket measurements to engine room intake measurements between 1939 and the end of 1941 (refs 6, 9). The corrections were made by adjusting the data before 1941 so that they are compatible with the data obtained from the mix of measurements during the 1961–90 period, and they

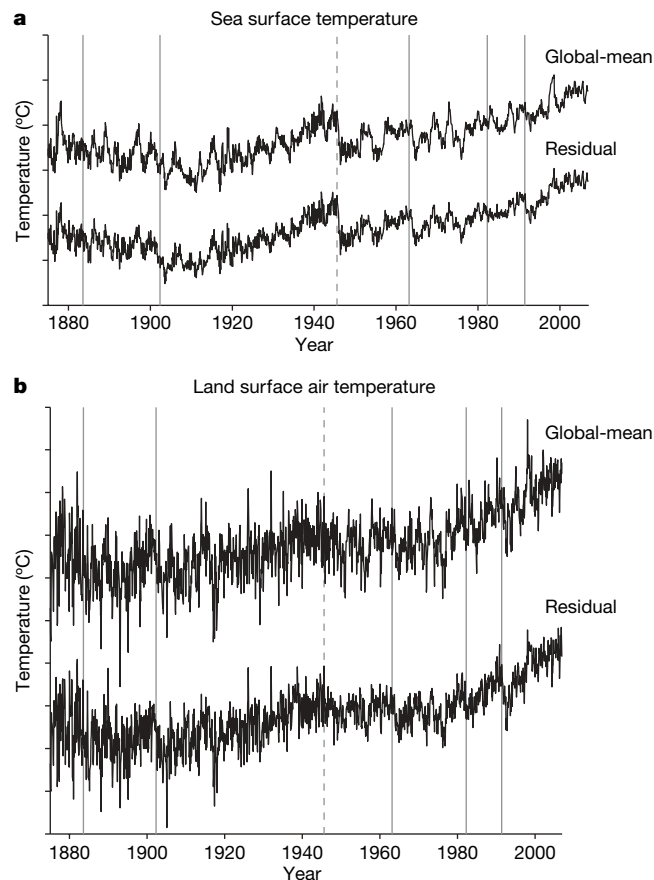


Figure 3 | Filtering the SST and land temperature time series. **a**, The original and ENSO-residual global-mean SST time series based on the HadSST2 data. **b**, The original and ENSO+COWL residual land surface temperature time series based on the CRUTEM3 land temperature data. Vertical lines are as in Fig. 2. Vertical axis shows temperature anomalies; tickmarks indicate steps of 0.5°C in both **a** and **b**. The residual global-mean SST time series is formed by subtracting T_{ENSO} from the unfiltered SST time series (the COWL index time series is not significantly correlated with the global-mean SST time series; Table 1).

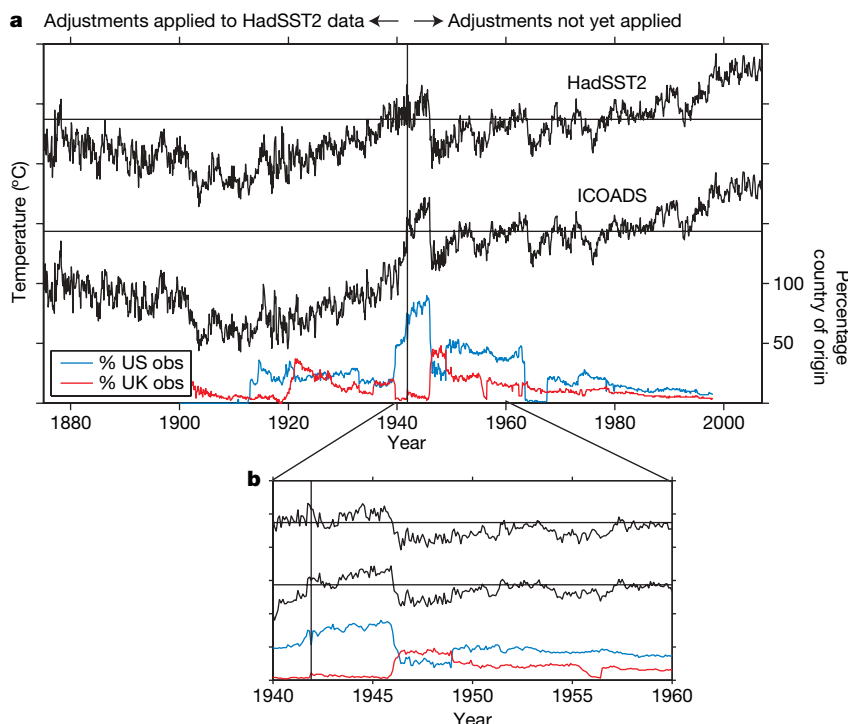


Figure 4 | The HadSST2, ICOADS and country-of-origin time series. **a**, Top, the ENSO residual global-mean SST time series reproduced from Fig. 3. Middle, as in the top time series but for data from ICOADS. Bottom, the percentage of observations which can be positively identified as coming from US (blue) and UK (red) ships. The vertical line denotes December 1941. Data to the left of the vertical line were corrected in the Met Office Hadley Centre

data^{6,9}; data to the right of the vertical line have not been adjusted in the HadSST2 data. **b**, As **a** but focused on the period 1940–60. Left vertical axis shows temperature anomalies; tickmarks indicate steps of 0.5 °C. Right vertical axis shows percentage of observations. Horizontal lines in the temperature time series denote the mean for 1961–90.

result in the large differences between the ICOADS and HadSST2 data before 1941 (Fig. 4). However, these are the only substantial corrections that have been applied to the HadSST2 data, in part because important information about the mix of measurements for the post-Second World War era has only recently been digitized¹¹. From the country-of-origin information in Fig. 4, it is clear that the SST archive—and hence the mix of measurement methods—continued to evolve considerably during the decades following 1941.

The most notable change in the SST archive following December 1941 occurred in August 1945. Between January 1942 and August 1945, ~80% of the observations are from ships of US origin and ~5% are from ships of UK origin; between late 1945 and 1949 only ~30% of the observations are of US origin and about 50% are of UK origin. The change in country of origin in August 1945 is important for two reasons: first, in August 1945 US ships relied mainly on engine room intake measurements whereas UK ships used primarily uninsulated bucket measurements¹², and second, engine room intake measurements are generally biased warm relative to uninsulated bucket measurements^{6,7}.

Hence, the sudden drop in SSTs in late 1945 is consistent with the rapid but uncorrected change from engine room intake measurements (US ships) to uninsulated bucket measurements (UK ships) at the end of the Second World War. As the drop derives from the composition of the ICOADS data set, it is present in all records of twentieth-century climate variability that include SST data.

The Met Office Hadley Centre is currently assessing the adjustments required to compensate for the step in 1945 and subsequent changes in the SST observing network. The adjustments immediately after 1945 are expected to be as large as those made to the pre-war data (~0.3 °C; Fig. 4), and smaller adjustments are likely to be required in SSTs through at least the mid-1960s, by which time the observing fleet was relatively diverse and less susceptible to changes in the data supply from a single country of origin⁹. The new adjustments

are likely to have a substantial impact on the historical record of global-mean surface temperatures through the middle part of the twentieth century. The adjustments are unlikely to significantly affect estimates of century-long trends in global-mean temperatures, as the data before ~1940 and after the mid-1960s are not expected to require further corrections for changes from uninsulated bucket to engine room intake measurements. However, compensation for a different potential source of bias in SST data in the past decade—the transition from ship- to buoy-derived SSTs—might increase the century-long trends by raising recent SSTs as much as ~0.1 °C, as buoy-derived SSTs are biased cool relative to ship measurements¹⁰.

METHODS SUMMARY

The data used in this study are the land-surface (CRUTEM3), sea-surface (HadSST2) and combined land- and sea-surface temperature (HadCRUT3) data obtained from the Climatic Research Unit (<http://www.cru.uea.ac.uk>) and described in refs 9 and 13. The methodologies used to form the ENSO and COWL index time series, and to filter out the ENSO and COWL index time series from the global-mean time series, are described in the Methods section.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 28 January; accepted 4 April 2008.

1. Trenberth, K. *et al.* in *Climate Change 2007: The Physical Science Basis* (eds Solomon, S. *et al.*) Ch. 3 (Cambridge Univ. Press, Cambridge, UK, 2007).
2. Hegerl, G. *et al.* in *Climate Change 2007: The Physical Science Basis* (eds Solomon, S. *et al.*) Ch. 9 (Cambridge Univ. Press, Cambridge, UK, 2007).
3. Yulaeva, E. & Wallace, J. M. The signature of ENSO in global temperature and precipitation fields derived from the microwave sounding unit. *J. Clim.* **7**, 1719–1736 (1994).
4. Wallace, J. M., Zhang, Y. & Renwick, J. A. Dynamic contribution to hemispheric mean temperature trends. *Science* **270**, 780–783 (1995).
5. Robock, A. *et al.* Climatic consequences of regional nuclear conflicts. *Atmos. Chem. Phys.* **7**, 2003–2012 (2007).

6. Folland, C. K. & Parker, D. E. Correction of instrumental biases in historical sea surface temperature data. *Q. J. R. Meteorol. Soc.* **121**, 319–367 (1995).
7. Kent, E. C. & Taylor, P. K. Toward estimating climatic trends in SST. Part I: Methods of measurement. *J. Atmos. Ocean. Technol.* **23**, 464–475 (2006).
8. Smith, T. M. & Reynolds, R. W. Bias corrections for historical sea surface temperatures based on marine air temperatures. *J. Clim.* **15**, 73–87 (2002).
9. Rayner, N. A. *et al.* Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: The HadSST2 data set. *J. Clim.* **19**, 446–469 (2006).
10. Worley, S. J., Woodruff, S. D., Reynolds, R. W., Lubker, S. J. & Lott, N. ICOADS Release 2.1 data and products. *Int. J. Climatol.* **25**, 823–842 (2005).
11. Kent, E. C., Woodruff, S. D. & Berry, D. I. Metadata from WMO Publication No. 47 and an assessment of voluntary observing ships observation heights in ICOADS. *J. Atmos. Ocean. Technol.* **24**, 214–234 (2007).
12. *Marine Observer's Handbook* 8th edn (HMSO, London, 1963).
13. Brohan, P., Kennedy, J. J., Harris, I., Tett, S. F. B. & Jones, P. D. Uncertainty estimates in regional and global observed temperature changes: A new dataset from 1850. *J. Geophys. Res.* **111**, D12106, doi:10.1029/2005JD006548 (2006).

Acknowledgements We thank D. Parker and S. Solomon for comments on the manuscript, and R. Reynolds for reviews. D.W.J.T. and J.M.W. were supported by the NSF Climate Dynamics Program under budget numbers ATM-0132190 and ATM-0613082 (D.W.J.T.) and ATM-0318675 (J.M.W.). J.J.K. was supported by the Joint Defra and MoD Programme, GA01101 (Defra) and CBC/2B/0417_Annex C5 (MoD). P.D.J. was supported by the US Department of Energy (DE-FG02-98ER62601).

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.W.J.T. (davet@atmos.colostate.edu).

METHODS

All results and analyses are based on monthly-mean data. Following ref. 3, the tropical-mean surface temperature response to ENSO is modelled as:

$$C \frac{d}{dt} T_{\text{ENSO}}(t) = F(t) - \beta T_{\text{ENSO}}(t) \quad (1)$$

where $T_{\text{ENSO}}(t)$ denotes the simulated response of monthly-mean tropical-mean surface temperature anomalies to ENSO variability; $F(t)$ is the anomalous flux of sensible and latent heat in the eastern tropical Pacific, parameterized as the average SST anomaly in the dynamically active cold-tongue region (5°N – 5°S , 180°W – 90°W) multiplied by (1) the fractional area of the tropics covered by the cold-tongue region ($\sim 5\%$), and (2) a coupling coefficient of $25 \text{ W m}^{-2} \text{ K}^{-1}$; β is a linear damping coefficient found by linearizing the Stefan–Boltzmann law, $\beta = 4\sigma T_{\text{E}}^3$, where σ is the Stefan–Boltzmann constant and the mean temperature of the tropical atmosphere, T_{E} is assumed to be $\sim 255 \text{ K}$; C is the heat capacity of the tropics per unit area and is determined empirically such that the correlation between $T_{\text{ENSO}}(t)$ and tropical-mean surface temperature anomalies is maximized (the resulting heat capacity corresponds to the entire atmosphere and the top $\sim 8 \text{ m}$ of the tropical ocean; $C = 5.2 \times 10^7 \text{ J m}^{-2} \text{ K}^{-1}$). The model was initialized with anomalies in the cold-tongue region starting at 1870 and the output T_{ENSO} was used from 1875 to 1995.

Following ref. 4, the COWL pattern loadings are found as:

$$\text{COWL}(x) = [T(x, t) - T_{\text{NH}}(t)] \cdot T_{\text{NH}}(t) \quad (2)$$

where $T(x, t)$ are Northern Hemisphere (30°N – 90°N) temperature anomalies given as a function of space, x , and time, t ; $T_{\text{NH}}(t)$ denotes the time series of Northern Hemisphere (30°N – 90°N) mean surface temperature anomalies, and the overbar denotes the time mean. The time series of the COWL pattern (T_{COWL}) is found by projecting $[T(x, t) - T_{\text{NH}}(t)]$ onto $\text{COWL}(x)$ at each time step.

The T_{ENSO} and T_{COWL} index time series were fitted to global-mean temperatures as:

$$x_{\text{fitted}}(t) = \frac{\overline{x'(t) \cdot T_G'(t)}}{x'^2(t)} \cdot x(t)$$

where $x(t)$ corresponds to the monthly-mean T_{ENSO} or T_{COWL} index time series; the fractional term on the right-hand side denotes the regression of $T_G(t)$ onto $x(t)$; and x_{fitted} corresponds to the component of $T_G(t)$ that is linearly congruent with variations in T_{ENSO} or T_{COWL} . The regression coefficients are based on detrended values of the data so that shared trends in the time series do not contribute to the linear fits (the data are detrended for the 1950–2006 period), and are calculated for the period 1950–2006 when the data coverage is best (the results are not sensitive to reasonable changes in the period of analyses and are virtually unchanged if the data are not detrended). The regressions were done separately for T_{ENSO} and T_{COWL} as the time series are effectively uncorrelated; virtually identical results were obtained for multiple regression. In Fig. 1, the fit to the T_{ENSO} and T_{COWL} index time series was computed for global-mean combined ocean and land temperature data; in Figs 3 and 4, the fits were computed for global-mean land and ocean temperature data separately.

LETTERS

Live birth in the Devonian period

John A. Long^{1,2,3}, Kate Trinajstić⁴, Gavin C. Young² & Tim Senden⁵

The extinct placoderm fishes were the dominant group of vertebrates throughout the Middle Palaeozoic era¹, yet controversy about their relationships within the gnathostomes (jawed vertebrates) is partly due to different interpretations of their reproductive biology^{2–5}. Here we document the oldest record of a live-bearing vertebrate in a new ptyctodontid placoderm, *Materpiscis attenboroughi* gen. et sp. nov., from the Late Devonian Gogo Formation of Australia (approximately 380 million years ago)⁶. The new specimen, remarkably preserved in three dimensions, contains a single, intra-uterine embryo connected by a permineralized umbilical cord. An amorphous crystalline mass near the umbilical cord possibly represents the recrystallized yolk sac. Another ptyctodont from the Gogo Formation, *Austroptyctodus gardineri*⁷, also shows three small embryos inside it in the same position. Ptyctodontids have already provided the oldest definite evidence for vertebrate copulation⁸, and the new specimens confirm that some placoderms had a remarkably advanced reproductive biology, comparable to that of some modern sharks and rays. The new discovery points to internal fertilization and viviparity in vertebrates as originating earliest within placoderms.

Placodermi McCoy, 1848

Ptyctodontida Gross, 1932

Materpiscis attenboroughi gen. et sp. nov.

Etymology. Generic name from the Latin meaning ‘mother fish’; species name in honour of Sir David Attenborough, who first drew attention to the Gogo fish sites in his 1979 series *Life on Earth*.

Holotype. WAM 07.12.1 (Western Australian Museum, Perth).

Age and locality. From the Stomatoporoid camp locality, Gogo Station, near Fitzroy Crossing, Western Australia (Late Devonian, early Frasnian).

Diagnosis. A small aspinothoracid ptyctodontid fish having an anteriorly pointed nuchal plate that participates in the posterior margin of the skull roof, broad roughly triangular-shaped preorbitals that meet mesially; the marginal plate has a large postorbital region with parallel rows of tubercles adorning it; the submarginal is strap-like and strongly curved mesially; robust triturating tooth plates that meet only at anterior tips, superognathals with moderately high anterior dorsal process. The body is scaleless.

Description. *Materpiscis* gen. nov. (Figs 1–3) is readily distinguished from all other known ptyctodontids by the combination of lacking a spinal plate with anteriorly pointed nuchal plate (Fig. 1e), and a much-expanded postorbital region on the marginal plate (Fig. 1f). The nuchal plate participates in the posterior margin of the skull roof and the centrals are unusually broad (Fig. 1e). The skull roof is unusual in having prominent raised ridges for the main sensory-line canals (Fig. 1e). The tooth plates can be moved against each other as in life. This shows that only the anterior-most biting edges of the jaws met in full occlusion (Fig. 1i). The trunk shield plates do not differ significantly from those of many other aspinothoracid

ptyctodontids, in particular *Austroptyctodus*⁷, except for proportions and ornamentation, the median dorsal having a low median crest. A restoration of the dermal armour is shown in Fig. 1h. A detailed description of the dermal and axial skeleton, braincase and dentition of *Materpiscis* gen. nov., and its embryonic skeletal morphology, will be published elsewhere.

The Gogo Formation has yielded the earliest phosphatized gnathostome muscle tissues and nerve fibres⁹, and the new specimen shows additional soft-tissue preservation never before recorded in any fossil. The new specimen is exceptional in revealing a small partial skeleton located within the upper body cavity of a pregnant, adult female ptyctodontid placoderm (Fig. 1a–d). The relatively complete adult skeleton includes much of the tail (Fig. 1a). The single, partially ossified juvenile skeleton inside the body cavity of the adult clearly belongs to the same species based on characteristic tooth-plate and marginal plate morphology (Fig. 1c, d). The small internal individual must be an embryo, rather than an ingested prey item, because the delicate bones show no breakage or etching from stomach acids, and the dentition is of similar gross morphology to that of the adult. The location of the preserved embryo close to the vertebral column suggests that it is in the uterus rather than the gut, a view supported by the presence of mineralized soft tissue forming a placental connection between embryo and mother. Furthermore, a second specimen from Gogo, previously described as *Austroptyctodus gardineri*⁷, has been re-examined and we have identified three small embryos in the same position posterior to the adult anterior lateral plate as in *Materpiscis* gen. nov. (Fig. 1g). Cases demonstrating viviparity in the fossil record are extremely rare, the best being Mesozoic ichthyosaurs^{10,11} and mosasauroids¹². Presumed fetal specimens of a holocephalan fish from the Lower Carboniferous of Montana¹³ may extend the record of viviparity back into the Palaeozoic era, but were not found in or near the body cavity of the mother. Our new example extends the definite record of vertebrate viviparity back by some 200 million years.

The embryos of *Materpiscis* gen. nov. and *Austroptyctodus* are about 25% of the adult size, based on proportions of the anterior lateral plate. The embryonic bones are remarkably thin and fragile, yet perfectly preserved, with the upper and lower dentition still in occlusion in *Materpiscis* gen. nov. (Figs 1b–d and 2g). The skull bones of this specimen are oriented posteriorly, indicating that the embryo was curled. Only some dermal plates are represented in the embryos. Absence of certain bones at an early developmental stage is also recorded in tiny (14–16 mm) juvenile skeletons of the antiarch placoderm *Asterolepis*¹⁴. Identified embryonic bones comprise paired preorbitals, a paranuchal, anterior dorsolateral, anterior lateral and a marginal plate, and both sets of tooth plates in articulation (Figs 1b–d, g and 2g, f). The supragnathals of *Materpiscis* gen. nov. meet at a distinct medially directed symphysis (sym, Fig. 2g), lending increased structural support for biting ability immediately after birth.

¹Museum Victoria, Melbourne, PO Box 666, Melbourne 3001, Australia. ²Research School of Earth Sciences, The Australian National University, Canberra 0200, Australia. ³School of Geosciences, Monash University, Clayton, Victoria 3800, Australia. ⁴School of Earth and Geographical Sciences, The University of Western Australia, Perth 6009, Western Australia, Australia. ⁵Department of Applied Mathematics, Research School of Physical Sciences and Engineering, The Australian National University, Canberra 0200, Australia.

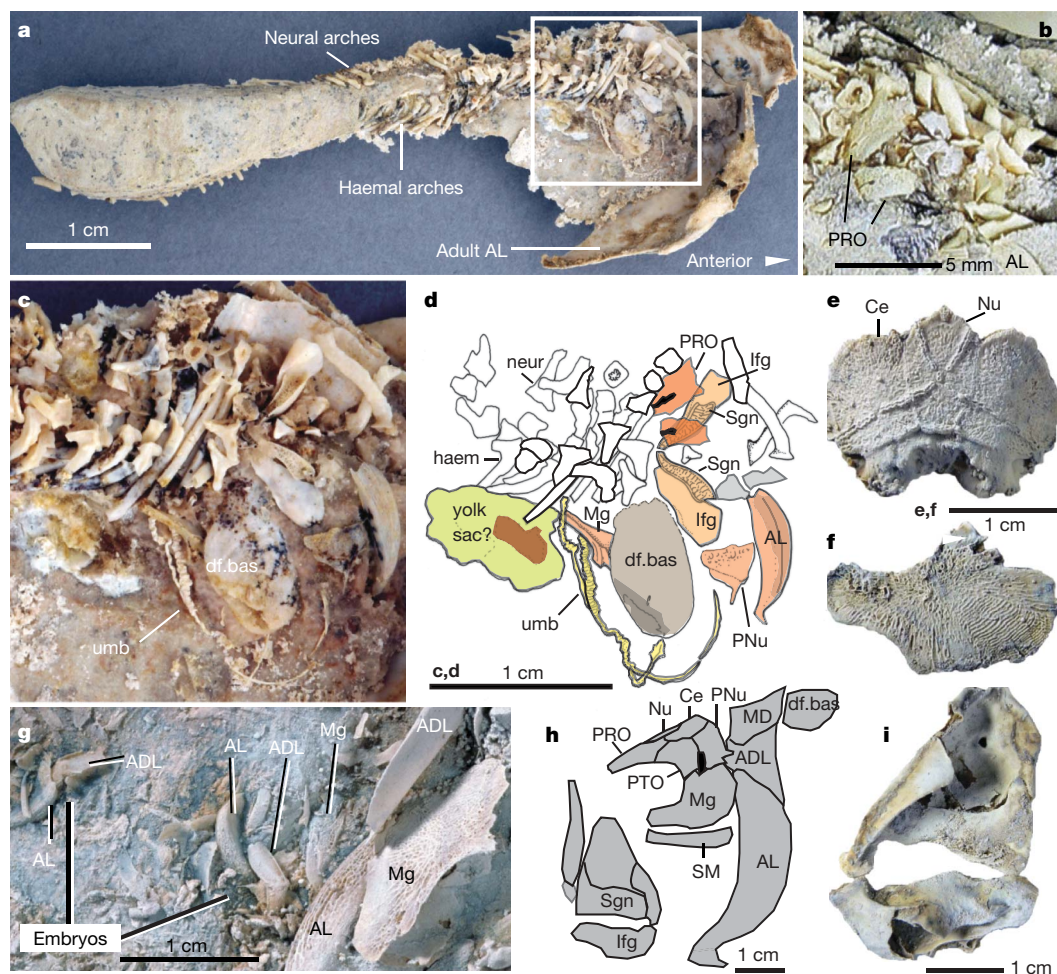


Figure 1 | Holotype of *Materpiscis attenboroughi* gen. et sp. nov. (WAM 07.12.1). **a**, Body with embryo. The boxed area is magnified in panel **c**. **b**, Embryo at early preparation. **c**, **d**, Umbilical cord and embryo at later preparation. **e**, Adult skull roof in dorsal view. **f**, Adult left marginal. **g**, Dermal armour restored in lateral view. **h**, Left adult tooth plates in mesial view. **i**, *Austroptyctodus gardineri* (WAM 86.9.662) showing embryos.

ADL, anterior dorsolateral plate; AL, anterior lateral plate; Ce, central; df.bas, dorsal fin basal plate of adult; haem, haemal arch; Ifg, inferognathal; MD, median dorsal plate; Mg, marginal; neur, neural arch; Nu, nuchal; PNU, paranuchal plate; PRO, preorbital; PTO, postorbital; Sgn, supragathal; SM, submarginal plate; umb, umbilical cord.

A thin tube of mineralized white material with slight helical twisting (umb, Fig. 1c, d) connects the *Materpiscis* gen. nov. embryo to a region within the upper body cavity of the female, marked by an

ovoid mass of coarsely crystalline calcite, possibly the position of the yolk sac (Fig. 1c, d). In modern sharks, the umbilical cord wall comprises multiple small vessels, capillaries and a dense spongy

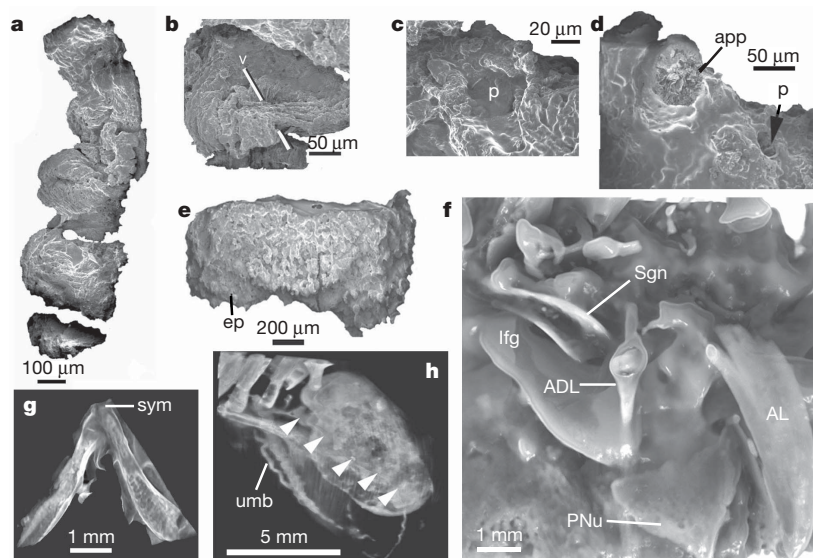


Figure 2 | Holotype of *Materpiscis attenboroughi* gen. et sp. nov. (WAM 07.12.1). **a–e**, SEM images of umbilical cord. **a**, Umbilicus section. **b**, Image showing a vascular canal (v). **c**, image showing pit (p). **d**, Image showing possible appendiculae (app) and pit (p). **e**, Image showing outer epithelium surface (ep) and stalk wall. **f**, High-resolution computer tomogram (HRCT) showing embryonic bones. **g**, HRCT tomogram showing supragathals in occlusion. **h**, HRCT showing umbilicus (white arrowheads). Abbreviations as for Fig. 1 except for: ep, epithelium; sym, symphysis.

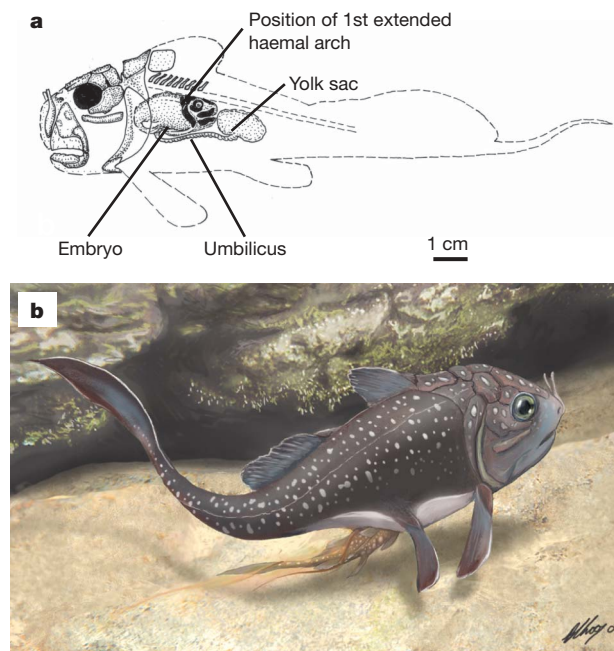


Figure 3 | *Materpiscis* gen. nov. **a**, Diagram showing position of embryo and yolk sac within the mother. **b**, Artist's reconstruction of *Materpiscis* gen. nov. giving birth (by B. Choo).

connective tissue¹⁵. The fossilized tube has similar external shape and position as the umbilical cord in chondrichthyans. High-resolution scanning electron microscopy (SEM) imaging and high-resolution computer tomography scanning of this specimen revealed that the cord loops underneath the adult dorsal fin basal plate (Fig. 2h and Supplementary Information). Its surface is covered with a thin layer of epithelium (Fig. 2e) and has a series of pits of varying sizes (Fig. 2c, d), with the wall of the umbilical cord made up of a series of vessels and a dense matrix of tissue. The path of one major vessel could be determined running through the cord (Fig. 2b). Branching off from the umbilical cord are the remains of filamentous processes equivalent to the appendiculae in extant sharks (app, Fig. 2d). Thus, ptyctodonts apparently evolved beyond the first live-bearing stage (oviviparity, in which the mother retains her eggs inside her body, providing only protection, with the yolk sac providing all nutrition) to matrotrophy. The relative size and position of the embryo and yolk sac in relation to the mother fish is shown in Fig. 3a, with an attempted restoration of the fish giving birth shown in Fig. 3b, based on the observation that several viviparous sharks give birth tail-first to their young. Further discussion of the significance of viviparity and evolution can be found in the Supplementary Information.

METHODS SUMMARY

Preparation followed the standard acid etching technique using dilute (10%) acetic acid, although fine-resolution acid bathing for short time intervals was used for revealing the very delicate bones of the embryo.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 10 January; accepted 3 April 2008.

- Denison, R. *Handbook of Paleichthyology. Placodermi* 128 (Gustav Fischer, Stuttgart, 1978).
- Gardiner, B. G. The relationships of placoderms. *J. Vert. Paleo.* **4**, 379–395 (1984).
- Young, G. C. The relationships of placoderm fishes. *Zool. J. Linn. Soc.* **88**, 1–57 (1986).
- Goujet, D. & Young, G. C. in *Recent Advances in the Origin and Early Radiation of Vertebrates* (eds Arratia, G., Wilson, M. & Cloutier, R.) 109–126 (Dr Freiderich Pfeil, Munich, 2004).
- Miles, R. S. & Young, G. C. in *Problems in Vertebrate Evolution* (eds Andrews, S. M., Miles, R. S. & Walker, A. D.) 123–198 (Linnean Soc. Symp. Series 4, 1977).
- Long, J. A. *Swimming in Stone—the Amazing Gogo Fossils of the Kimberley* 320 (Fremantle Arts Centre, Perth, 2006).
- Long, J. A. Ptyctodontid fishes (Vertebrata, Placodermi) from the Late Devonian Gogo Formation, Western Australia, with a revision of the European genus *Ctenurella* Ørvig, 1960. *Geodiversitas* **19**, 515–555 (1997).
- Miles, R. S. Observations on the ptyctodont fish, *Rhamphodopsis* Watson. *J. Linn. Soc. Zool.* **47**, 99–120 (1967).
- Trinajstić, K., Marshall, E., Long, J. & Bifield, K. Exceptional preservation of nerve and muscle tissues in Devonian placoderm fish and their phylogenetic implications. *Biol. Lett.* **3**, 197–200 (2007).
- Böttcher, R. Neue Erkenntnisse über die Fortpflanzungsbiologie der Ichthyosaurier (Reptilia). *Stuttgarter Beitr. Nat.* **B 164**, 1–51 (1990).
- Maxwell, E. E. & Caldwell, M. W. First record of live birth in Cretaceous ichthyosaurs: closing an 80 million year gap. *Proc. R. Soc. Lond. B* **270** (Suppl. 1), S104–S107 (2003).
- Caldwell, M. W. & Lee, M. S. Y. Live birth in marine lizards (mosasauroids). *Proc. R. Soc. Lond. B* **268**, 2397–2401 (2001).
- Lund, R. Viviparity and interuterine feeding in a new holocephalan fish from the Lower Carboniferous of Montana. *Science* **209**, 697–699 (1980).
- Upeniec, I. The unique fossil assemblage from the Lode Quarry (Upper Devonian, Latvia). *Mitt. Mus. Natkd. Berl. Geowiss.* **4**, 101–119 (2001).
- Wourms, J. Maximization of evolutionary trends for placental viviparity in the spadenose shark, *Scoliodon laticaudus*. *Env. Biol. Fish.* **38**, 269–294 (1993).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Gomon and M. Lee for discussion of the material, and M. Caldwell for comments on the paper. P. Lillywhite assisted with photography. J.A.L., K.T. and G.C.Y. are supported by an Australian Research Council Discovery Grant. We thank L. Hatcher for finding the specimen on the 2005 Museum Victoria Gogo Expedition. The specimen was partially prepared by D. Pickering.

Author Contributions The fine preparation of the new specimen was done by J.A.L., and it was described by J.A.L., K.T. and G.C.Y.; T.S. participated in the 2005 Gogo expedition and analysed the specimen using XCT scan imagery.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.A.L. (jlong@museum.vic.gov.au).

METHODS

Specimen preparation. The acid immersion procedure (in 10% acetic acid for 2 days, washed in running water for a day, air-dried, all exposed bones then hardened with Mowital B30 consolidant dissolved in pure ethanol) was repeated over 2 months to extract fully the adult dermal plates of the head and trunk shields. The first part of the initial preparation was carried out by D. Pickering. Once enough of the skeleton was revealed to show that it was a ptyctodontid placoderm, the specimen was handed over to J.A.L. who continued preparation, taking photographs of the specimen after every acid bath cycle to record exact positions of individual skeletal elements.

The remaining elongate section of rock contained what we originally thought was just the tail vertebrae and parts of the axial skeleton, so preparation was halted to retain a high degree of articulation in the tail for study of the articulated vertebral units. However, during a visit to Museum Victoria in November 2007 by K.T., it was decided to prepare the tail section of the specimen one more time to clarify details of the anterior axial skeleton, so the specimen was placed in weak acetic acid (7%) for one night. The following day the embryonic bones were first revealed. From that point on we decided to take much more caution with preparation. We noted from the previous photographs that some delicate embryonic bones were already exposed previously and missed by us, dismissed as possibly being thin body scales as occur on the flank of *Campbellodus*⁷. After the acid bath the delicate preorbital plates of the embryo were now missing, but in the process the rest of the embryonic skeleton, and the umbilical cord and yolk sac, had been revealed. A search of the kept residue from the last acid bath revealed the missing embryonic preorbital plates, which were so fragile that one was severely damaged trying to lift it onto a slide mount.

One more attempt was made to prepare the specimen further to reveal more of the embryonic skeleton. This time the specimen was placed in very dilute acetic acid (5%) for 1 h at each immersion. The wet specimen was then carefully examined under a microscope and re-immersed in acid until the desired level of detail was exposed, after three more immersions, as is shown in the illustrations of our paper. During this time the thin layer of limestone holding the articulate tail together dissolved through, dividing the specimen into two sections: one containing embryo, umbilicus and yolk sac with parts of the adult axial skeleton and dermal trunk shield, and a smaller elongate section containing the distal axial skeleton.

Tomogram of specimen. High-resolution X-ray tomography was conducted in-house to explore the relationship between external and internal structures at a spatial resolution of 13 μm ^{16,17}. The included animation (Supplementary Movie) shows the specimen rotating to reveal parts of the umbilicus not visible on the surface, which are looped underneath a flat-lying dorsal fin basal plate from the mother fish.

All volume rendering and animation was done using the freeware package, Drishti (<http://anuf.anu.edu.au/Vizlab/drishti>).

Scanning electron microscopy. Scanning electron micrographs were taken at The Centre for Microscopy, Characterization and Analysis, University of Western Australia. Small portions of the umbilicus were removed from the specimen and imaged using a Zeiss 1555 VP-FESEM. Samples were first coated in carbon and then gold. SEM was chosen as it gave finer resolution for external surface detail than the XCT scans.

16. Sakellariou, A., Sawkins, T. J., Senden, T. J. & Limaye, A. X-ray tomography for mesoscale physics applications. *Physica A* **339**, 152–158 (2004).
17. Sakellariou, A. *et al.* An x-ray tomography facility for quantitative predictions of mechanical and transport properties in geological, biological and synthetic systems. In *Development in X-Ray Tomography IV* (ed. Bonse, U.) *Proc. SPIE Vol.* 5535 473–484 (SPIE, Bellingham, Western Australia, 2004).

Abundance and diversity of microbial life in ocean crust

Cara M. Santelli^{1,2}, Beth N. Orcutt³, Erin Banning¹, Wolfgang Bach^{2,4}, Craig L. Moyer⁵, Mitchell L. Sogin⁶, Hubert Staudigel⁷ & Katrina J. Edwards^{2,3}

Oceanic lithosphere exposed at the sea floor undergoes seawater–rock alteration reactions involving the oxidation and hydration of glassy basalt. Basalt alteration reactions are theoretically capable of supplying sufficient energy for chemolithoautotrophic growth¹. Such reactions have been shown to generate microbial biomass in the laboratory², but field-based support for the existence of microbes that are supported by basalt alteration is lacking. Here, using quantitative polymerase chain reaction, *in situ* hybridization and microscopy, we demonstrate that prokaryotic cell abundances on seafloor-exposed basalts are 3–4 orders of magnitude greater than in overlying deep sea water. Phylogenetic analyses of basaltic lavas from the East Pacific Rise (9° N) and around Hawaii reveal that the basalt-hosted biosphere harbours high bacterial community richness and that community membership is shared between these sites. We hypothesize that alteration reactions fuel chemolithoautotrophic microorganisms, which constitute a trophic base of the basalt habitat, with important implications for deep-sea carbon cycling and chemical exchange between basalt and sea water.

We assessed the abundance, species richness and phylogenetic diversity of endolithic and epilithic microbial communities inhabiting young, unsedimented ocean crust at the sea floor. Basaltic seafloor lavas from the East Pacific Rise (EPR), 9° N (Fig. 1a; Supplementary Fig. 1a, b), were sampled to represent various ages (up to 20 thousand years) and alteration states (Fig. 1b–e; Supplementary Table 1) for analysis. Quantitative polymerase chain reaction (qPCR) measurements (Supplementary Table 2) of the glassy rinds of lava flows indicate total bacterial and archaeal cell densities range from $\sim 3 \times 10^6$ to $\sim 1 \times 10^9$ cells per g rock (10^7 to 3×10^9 gene copies per g rock). Bacteria dominated (88–96%) all samples examined. Catalysed reporter deposition–fluorescent *in situ* hybridization (CARD–FISH) analyses revealed dense populations of Bacteria on basalt surfaces, which were significantly more abundant than Archaea (Fig. 2a–c). In contrast, the total microbial cell number of deep ocean waters is only 8×10^3 cells per ml sea water to 9×10^4 cells per ml sea water^{3,4} in the deep ocean (>1,000 m), half of which are Archaea.

The bacterial community composition was evaluated using full-length 16S ribosomal RNA gene clone libraries that were constructed from basalt samples and surrounding sea water. Statistical approaches were applied to evaluate species richness, the number of operational taxonomic units (OTUs) of these communities, as compared to other environments. We define an OTU at 97% sequence similarity—a commonly recognized level for comparative analysis. Rarefaction was used to compare bacterial richness of the

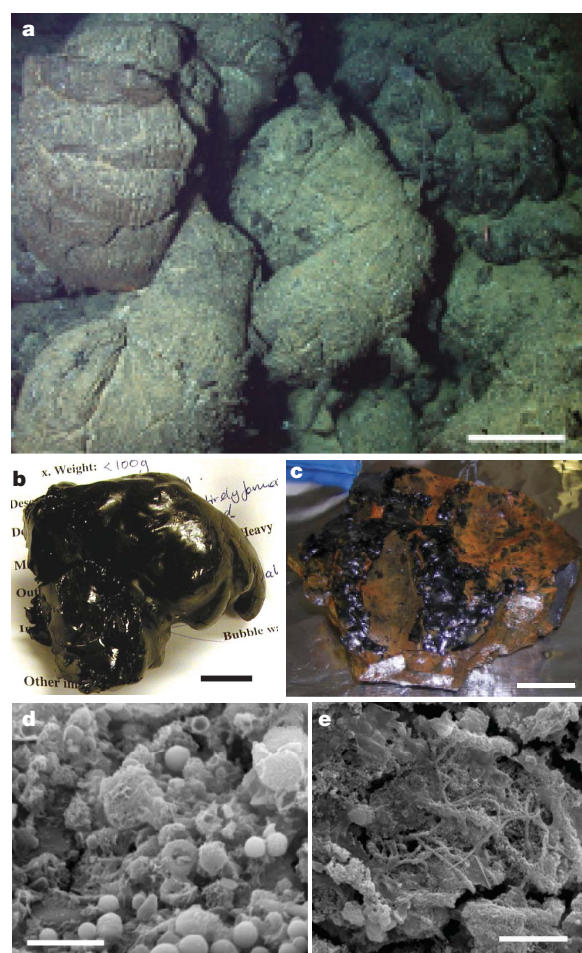


Figure 1 | A highly diverse and abundant epilithic and endolithic microbial community exists on basaltic lavas from the East Pacific Rise. **a**, The EPR at 9° N is characterized by lava flows, such as pillow basalts that outcrop at the sea floor. Scale bar, 40 cm. **b**, **c**, Photographs showing the range of volcanic samples used in this study from fresh and glassy (**b**) to more altered and oxide-coated (**c**). Scale bars, 2 cm (**b**) and 4 cm (**c**). **d**, **e**, Scanning electron microscopy (SEM) images of different presumed cellular morphologies, such as coccoidal (**d**) and filamentous (**e**) structures that were observed on ferromanganese-oxide-encrusted and iron-oxide-coated samples. Scale bars, 5 μ m.

¹MIT/WHOI Joint Program in Oceanography and Ocean Engineering and ²Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA. ³Geomicrobiology Group, Department of Biological Sciences, Marine Environmental Biology, University of Southern California, 3616 Trousdale Boulevard, Los Angeles, California 90089-0371, USA. ⁴Fachbereich Geowissenschaften, Universität Bremen, Postfach 33 04 40, D-28334 Bremen, Germany. ⁵Department of Biology, Western Washington University, Bellingham, Washington 98225, USA. ⁶Josephine Bay Paul Center, Marine Biological Laboratory at Woods Hole, 7 MBL Street, Woods Hole, Massachusetts 02543, USA. ⁷Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, La Jolla, California 92037, USA.

collective EPR basalts with that of other oceanic habitats (Fig. 3a) and other known high-diversity environments (Fig. 3b). Although richness varies among the samples (Fig. 3c), cumulative treatment is warranted here for comparison with the selective studies, which similarly sum multiple samples from a single habitat (see Supplementary Information). The comparative oceanic habitats included a hydrothermal white smoker from the EPR⁵, the upper water column of the Sargasso Sea⁶, hydrothermal vent fluids from the Mid-Atlantic Ridge (MAR) captured in an *in situ* growth chamber⁷, and deep-subsurface sediments from the Nankai Trough⁸. High-diversity environments included for comparison were terrestrial farm soil⁹ and a hypersaline microbial mat from the Guerrero Negro¹⁰. Although the environments were not sampled to saturation, rarefaction analysis revealed that EPR basalts harbour greater bacterial richness than the other marine habitats (Fig. 3a), and are most comparable to richness in mat and soil microbial communities (Fig. 3b).

To test whether the EPR basalt richness is site-specific or common among seafloor-exposed basalts, we also examined 16S rRNA gene clone libraries generated from lavas collected from offshore the big island of Hawaii (hereafter referred to as Hawaii; Supplementary Fig. 1a, c and Supplementary Table 1). The individual (Fig. 3c) and cumulative (Fig. 3b) richness of the Hawaii basalts as estimated by rarefaction analysis are similar to the EPR richness estimates.

Comparative rarefaction results were corroborated by nonparametric-based estimates of total richness and diversity (Supplementary Table 3). The Chao1 total richness estimate for the cumulative EPR basalts is 440 OTUs (lower and upper 95% confidence intervals of 338 and 609 OTUs). The Chao1 estimate for farm soil, the high-diversity

end-member, is 1,410 OTUs (95% confidence interval of 1,223–1,658 OTUs), whereas the EPR deep-seawater survey, the low-diversity end-member, is 12 OTUs (95% confidence interval of 11–22 OTUs). Because nonparametric estimators produce widely varying results

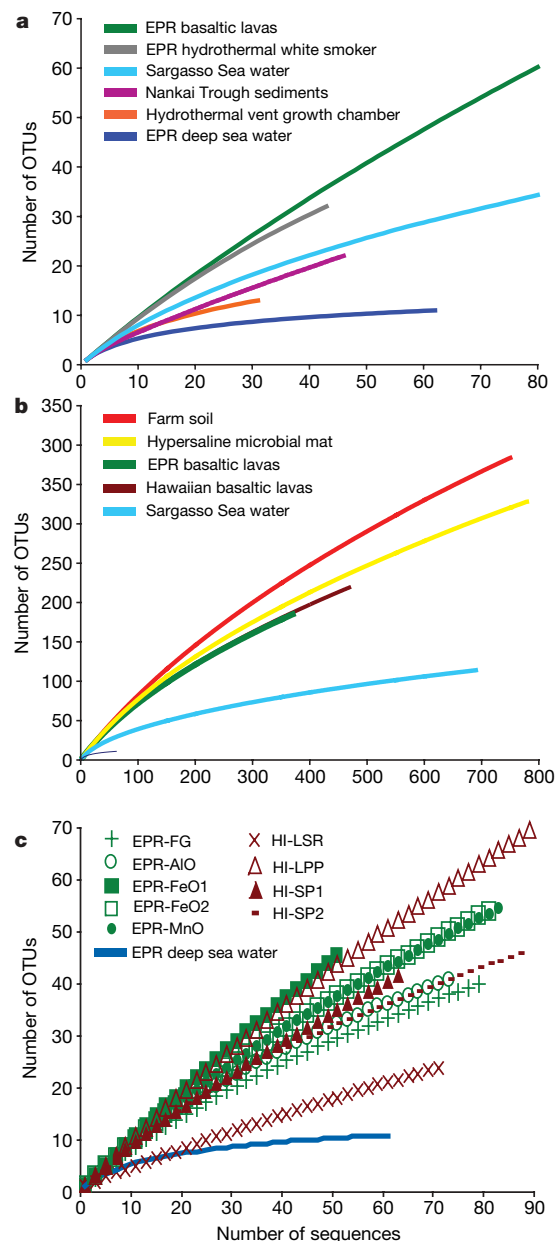


Figure 3 | Relative bacterial richness from several environmental studies shown through rarefaction analyses. **a**, Species richness of Bacteria inhabiting EPR seafloor lavas (cumulative results) is compared with that of other ocean environments, such as the Sargasso Sea⁶ (partial curve), a MAR hydrothermal vent *in situ* growth chamber⁷, an EPR hydrothermal white smoker spire⁵, Nankai Trough deep-sea sediments⁸, and EPR deep sea water. The bacterial richness of the EPR basalts is also compared to a basalt-hosted community from Hawaii and other known high-richness environments (**b**), such as a farm soil⁹ (partial curve) and a hypersaline microbial mat from the Guerrero Negro¹⁰ (partial curve). Partial rarefaction curves are shown for visualization purposes; however, complete data sets were used in calculating curve projections. **c**, Rarefaction curves for the individual EPR and Hawaii basalt clone libraries. A partial curve is shown for HI-LPP (total clones = 246). Comparative studies in **a** and **b** are based on near full-length 16S rRNA gene sequences, and most studies are the sum of several environmental samples. OTUs are defined at a sequence similarity of $\geq 97\%$. FG, fresh glass; LSR, Loihi seamount South Rift; LPP, Loihi seamount Pisces Peak; SP1 & SP2, South Point samples.

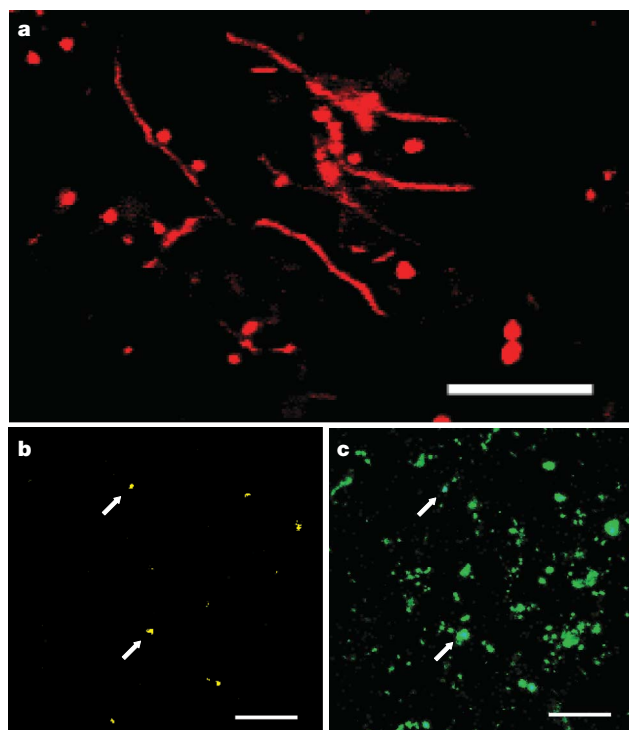


Figure 2 | Confocal laser scanning micrographs of cells hybridized with domain-specific fluorescent probes (via CARD-FISH) on lava surfaces. Microbial cells were hybridized with either probe EUB338(I-III) mix (**a**) or probe ARCH915 (**b**) to target Bacteria or Archaea, respectively. A variety of bacterial cell morphologies such as filaments, cocci and rods were confirmed with hybridizations (**a**). Side-by-side comparisons of Archaea (**b**) versus total cells (**c**) reveals that Archaea account for only a small portion of the total cells. Total cells were identified with the general DNA stain SYBR green. The arrows point to some cells that overlap in each frame. Scale bars, 10 μm (**a**) and 20 μm (**b**, **c**).

and underestimate diversity^{11,12}, we do not interpret these results as accurate predictions of true species richness. Rather, these indicators support our rarefaction observations of the relative differences between environments. This secondary comparison is informative because rarefaction curves can intersect and cross with increased sampling effort, particularly in high-diversity communities in which curves do not plateau¹³. Results showed that the relative estimates of diversity were consistent among the different environments, regardless of the estimator or index used. One caveat of all comparative measurements is that the specific methods used to generate sequences (that is, DNA extraction protocol, PCR amplification, clone library construction) differed slightly in each study, which could influence the observed and predicted phylogenetic diversity. Our conclusions concerning the diversity of basalts relative to other habitats, however, seem unlikely to be altered by these differences.

Examination of the taxonomic groups obtained from various environments (Supplementary Fig. 2) reveals that the phylogenetic diversity of Bacteria in seafloor basalts is distinct from the biotopes of the other habitats that were used for comparison. The 21 taxonomic groups recovered from basalt are dominated by sequences belonging to the Proteobacteria phylum (all subdivisions), which account for 68% and 66% of all sequences in the EPR and Hawaii basalt libraries, respectively (EPR/Hawaii convention is used hereafter). Non-Proteobacteria groups recovered include the Planctomycetes (8%/5%, EPR/Hawaii), Actinobacteria (7%/8%), Bacteroidetes (4%/1%), Acidobacteria (3%/4%) and Verrucomicrobia (2%/2%). The other phyla recovered represent <2% each of the total clone library, and a small portion (6%/1%) could not be assigned a taxonomic group. Notably, estimates of shared OTU richness for the two geographically separated basalt communities (Supplementary Fig. 3) show considerable overlap in community membership, suggesting that oceanic basalt microbes are widely distributed among this biotope.

The phylogenetic distribution and richness of Bacteria in the other deep-sea environments are lower than deep-sea basalts (Supplementary Fig. 2). The deep seawater library reveals three taxonomic groups, all Proteobacteria: γ -proteobacteria (89% of the total library), α -proteobacteria and β -proteobacteria. Other deep-seawater studies^{14,15} have recovered additional taxa belonging to δ -proteobacteria, Bacteroidetes and Planctomycetes. Rarefaction analysis including those taxa (see Supplementary Information and Supplementary Figs 4 and 5) does not, however, significantly increase species richness. To assess the influence of hydrothermal source proximity on phylogenetic diversity, we compare the basalt data with clone library data constructed from several mineralogical layers in a hydrothermal white smoker spire at the EPR⁵. Supplementary Fig. 2 shows that the hydrothermal spire harbours different and fewer taxonomic groups, and in particular prevalence of ϵ -Proteobacteria, Aquificales and Thermales (microbes typically found in hydrothermal environments¹⁶) which infrequently occur in the other deep-sea libraries (basalts and sea water).

The differences in phylogenetic diversity, species richness, and total biomass between the basaltic lavas and overlying sea water raise questions about what energy source(s) fuel this biosphere. Potential energy sources capable of sustaining microbial life in basaltic ocean crust include hydrothermal input of both dissolved and particulate manganese and iron phases to the seafloor environment that could provoke redox reactions capable of supporting chemolithoautotrophic growth^{1,17}. Reduction reactions with iron or manganese as the electron acceptors could occur if enough hydrogen is present and anaerobic conditions exist. Dissolved organic carbon in sea water or hydrothermal fluids¹⁸ could also provide energy to heterotrophic microorganisms. It is suggested, however, that most dissolved organic carbon in the deep sea is recalcitrant to microbial oxidation¹⁹, but this remains largely unexplored.

Alteration reactions occurring on basalt surfaces could support microbial growth¹. We inferred that such reactions contribute to

biomass production within the basalt-hosted biotope and examined the possibility from a theoretical perspective. Lava surfaces are composed predominantly of volcanic glass, a highly reactive rock component that contains reduced elemental species such as iron, sulphur and manganese. Oxygen and nitrate in deep sea water oxidize these reduced constituents, and chemolithoautotrophic microorganisms could potentially exploit the free energy changes associated with these redox reactions for their metabolic requirements¹.

Support for this hypothesis is provided by laboratory studies that demonstrate that chemolithoautotrophic iron-oxidizing bacteria isolated from the sea floor are able to use rock and minerals, including basalt glass, for metabolism and growth². Furthermore, biomass production calculations¹ suggest that the observed oxidation of iron and sulphur in the upper ocean crust, if entirely mediated by microorganisms, can support the fixation of up to $\sim 5 \times 10^{11}$ g C yr⁻¹. Assuming $\sim 4 \times 10^{15}$ g of fresh basalt yr⁻¹ are subject to oxidation reactions¹ and a single cell contains 200 fg C (ref. 20), we estimate that 6×10^7 cells per g rock may be supported through alteration reactions (if microbial growth efficiency is 0.1 and biomass turnover occurs once per year). These parameters can be adjusted for more conservative estimates preferring a more oligotrophic or slower growing environment in which more energy goes into cell maintenance rather than cell growth, as in deep-sea sediments²¹ (for example, 6×10^9 cells per g basalt could be produced if carbon content is 20 fg C cell⁻¹, growth efficiency is 0.01, and biomass turnover is 100 years). Cell densities in volcanic glasses from the EPR established through qPCR measurements, $\sim 10^6$ to 10^9 cells per g basalt, fall within the range of predicted densities of $\sim 10^7$ to 10^9 cells per g basalt based on steady state iron and sulphur oxidation rates.

Our calculations suggest that alteration reactions in the upper ocean crust may fuel microbial ecosystems at the sea floor and contribute to biomass production and diversity in these systems. This hypothesis supports the understanding of the phylogenetically rich and distinct nature of the basalt biotope, otherwise it may be anticipated that the basalt-hosted community would more closely match that of the surrounding environment(s), that is, sea water in the case of unsedimented mid-ocean ridges like the EPR. Enrichment of taxa from diverse metabolic groups may result from the establishment of micro-environments within or on rock cavities and surfaces during alteration, secondary mineral precipitation and biofilm formation. Niche creation would allow for a greater variety of redox reactions and metabolic pathways (for example, heterotrophic, anaerobic, or reductive) within small spatial scales, including those supporting organotrophic and mixotrophic communities. The observed microbial biomass and phylogenetic diversity may consequently be an expression of the range of diverse chemical microenvironments that develop during basalt alteration.

METHODS SUMMARY

Sample collection and analysis. Basaltic seafloor lavas were collected from the EPR at 9° N (Supplementary Fig. 1a, b), during R/V *Atlantis* research cruises AT11-7 and AT11-20 and around Hawaii (Supplementary Fig. 1a, c) during *Ka'imikai-o-Kanaloa* cruise KOK 02-24 using the submersibles *Alvin* and *Pisces V*, respectively. DNA was extracted using the Ultraclean soil DNA kit (MoBio Laboratories) with minor modifications. Near full-length 16S rRNA gene clone libraries were constructed as previously described²² and sequenced at the Josephine Bay Paul Center (JBPC) at the Marine Biological Laboratories (MBL).

Phylogenetic analysis. Sequences were assembled and edited using an automated pipeline at the JBPC and aligned with the ARB software package (<http://www.arb-home.de/>). Phylogenetic affiliation was assigned with 'Classifier' available through the Ribosomal Database Project²³ using an 80% confidence threshold. Rarefaction analysis and species richness estimates were performed using the program Distance Based OTU and Richness Determination (DOTUR)²⁴, and the program Shared OTUs and Similarity (SONS)²⁵ was used for shared community richness estimates. Sequences from the MAR hydrothermal vent *in situ* growth chamber⁷, from the Guerrero Negro hypersaline microbial mat¹⁰ and from the EPR hydrothermal white smoker spire⁵ were accessed from GenBank. The full Minnesota farm soil and Nankai Trough deep

subsurface sediment⁸ data sets were provided with permissions from the authors⁹. The Sargasso Sea⁶ data, previously analysed in ref. 24, were downloaded from the DOTUR website (<http://www.plantpath.wisc.edu/fac/joh/DOTUR.html>).

CARD–FISH and quantitative PCR analysis. For CARD–FISH analyses, para-formaldehyde-fixed samples were hybridized with horseradish-peroxidase-linked oligonucleotide probes EUB338-I, EUB338-II and EUB338-III (ref. 26) or ARCH915 (ref. 27) with Cy3-tyramides (Biomers.net). SYBR Green I DNA stain (Invitrogen) was used as a counterstain. Digital images were collected on a LSM510-META laser scanning confocal microscope (Carl Zeiss) at the MBL and processed with the LSM Image Browser version 3.5 and Adobe Photoshop. Environmental DNA from lavas was used for qPCR to quantify bacterial²⁸ and archaeal²⁹ 16S rDNA gene copies. Cell densities were approximated using a conversion of 4.1 or 1.6 copies cell for Bacteria or Archaea respectively³⁰.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 10 February; accepted 12 March 2008.

- Bach, W. & Edwards, K. J. Iron and sulfide oxidation within the basaltic ocean crust: Implications for chemolithoautotrophic microbial biomass production. *Geochim. Cosmochim. Acta* **67**, 3871–3887 (2003).
- Edwards, K. J., Rogers, D. R., Wirsén, C. O. & McCollom, T. M. Isolation and characterization of novel psychrophilic, neutrophilic, Fe-oxidizing, chemolithoautotrophic α - and γ -Proteobacteria from the deep sea. *Appl. Environ. Microbiol.* **69**, 2906–2913 (2003).
- Karner, M. B., DeLong, E. F. & Karl, D. M. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**, 507–510 (2001).
- Cowen, J. P. *et al.* Fluids from aging ocean crust that support microbial life. *Science* **299**, 120–123 (2003).
- Kormas, K. A., Tivey, M. K., Von Damm, K. & Teske, A. Bacterial and archaeal phylotypes associated with distinct mineralogical layers of a white smoker spire from a deep-sea hydrothermal vent site (9° N, East Pacific Rise). *Environ. Microbiol.* **8**, 909–920 (2006).
- Venter, J. C. *et al.* Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**, 66–74 (2004).
- Reysenbach, A.-L., Longnecker, K. & Kirshtein, J. Novel bacterial and archaeal lineages from an in situ growth chamber deployed at a mid-atlantic ridge hydrothermal vent. *Appl. Environ. Microbiol.* **66**, 3798–3806 (2000).
- Kormas, K. A., Smith, D. C., Edgcomb, V. & Teske, A. Molecular analysis of deep subsurface microbial communities in Nankai Trough sediments (ODP Leg 190, Site 1176). *FEMS Microbiol. Ecol.* **45**, 115–125 (2003).
- Tringe, S. G. *et al.* Comparative metagenomics of microbial communities. *Science* **308**, 554–557 (2005).
- Ley, R. E. *et al.* Unexpected diversity and complexity of the Guerrero Negro hypersaline microbial mat. *Appl. Environ. Microbiol.* **72**, 3685–3695 (2006).
- Hong, S.-H., Bunge, J., Jeon, S.-O. & Epstein, S. S. Predicting microbial species richness. *Proc. Natl Acad. Sci. USA* **103**, 117–122 (2006).
- Hughes, J., Hellmann, J., Ricketts, T. & Bohannan, B. Counting the uncountable: statistical approaches to estimating microbial diversity. *Environ. Microbiol.* **67**, 4399–4406 (2001).
- Hughes, J. B. & Hellmann, J. J. The application of rarefaction techniques to molecular inventories of microbial diversity. *Methods Enzymol.* **397**, 292–308 (2005).
- Huber, J. A., Butterfield, D. A. & Baross, J. A. Bacterial diversity in a subseafloor habitat following a deep-sea volcanic eruption. *FEMS Microbiol. Ecol.* **43**, 393–409 (2003).
- Huber, J. A., Johnson, H. P., Butterfield, D. A. & Baross, J. A. Microbial life in ridge flank crustal fluids. *Environ. Microbiol.* **8**, 88–99 (2006).
- Huber, J. A. *et al.* Microbial population structure in the deep marine biosphere. *Science* **318**, 97–100 (2007).
- Edwards, K. J., Bach, W. & McCollom, T. M. Geomicrobiology in oceanography: mineral–microbe interactions at and below the seafloor. *Trends Microbiol.* **13**, 449–459 (2005).
- Lang, S. Q., Butterfield, D. A., Lilley, M. D., Johnson, H. P. & Hedges, J. I. Dissolved organic carbon in ridge-axis and ridge-flank hydrothermal systems. *Geochim. Cosmochim. Acta* **70**, 3830–3842 (2006).
- Barber, R. T. Dissolved organic carbon from deep waters resists microbial oxidation. *Nature* **220**, 274–275 (1968).
- Whitman, W. B., Coleman, D. C. & Wiebe, W. J. Prokaryotes: The unseen majority. *Proc. Natl Acad. Sci. USA* **95**, 6578–6583 (1998).
- Biddle, J. F. *et al.* Heterotrophic archaea dominate sedimentary subsurface ecosystems of Peru. *Proc. Natl Acad. Sci. USA* **103**, 3846–3851 (2006).
- Rogers, D. R., Santelli, C. M. & Edwards, K. J. Geomicrobiology of deep-sea deposits: estimating community diversity from low-temperature seafloor rocks and minerals. *Geobiology* **1**, 109–117 (2003).
- Cole, J. R. *et al.* The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Res.* **31**, 442–443 (2003).
- Schloss, P. D. & Handelsman, J. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl. Environ. Microbiol.* **71**, 1501–1506 (2005).
- Schloss, P. D. & Handelsman, J. Introducing SONS, a tool for OTU-based comparisons of membership and structure between microbial communities. *Appl. Environ. Microbiol.* **72**, 6773–6779 (2006).
- Daims, H., Bruhl, A., Amann, R., Schleifer, K.-H. & Wagner, M. Probe EUB338 is insufficient for the detection of all bacteria: development and evaluation of a more comprehensive probe set. *Syst. Appl. Microbiol.* **22**, 438–448 (1999).
- Raskin, L., Stromley, J., Rittmann, B. E. & Stahl, D. A. Group-specific 16S rRNA hybridization probes to describe natural communities of methanogens. *Appl. Environ. Microbiol.* **60**, 1232–1240 (1994).
- Nadkarni, M., Martin, F. E., Jacques, N. A. & Hunter, N. Determination of bacterial load by real-time PCR using a broad range (universal) probe and primer set. *Microbiology* **148**, 257–266 (2002).
- Takai, K. & Horikoshi, K. Rapid detection and quantification of members of the archaeal community by quantitative PCR using fluorogenic probes. *Appl. Environ. Microbiol.* **66**, 5066–5072 (2000).
- Klappenbach, J. A., Saxman, P. R., Cole, J. R. & Schmidt, T. M. rrndb: the ribosomal RNA operon copy number database. *Nucleic Acids Res.* **29**, 181–184 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The authors thank D. Fornari, M. Tivey and H. Schouten for allowing C.M.S. and W.B. to participate in their research cruise AT11-7 to collect samples at the EPR, D. Rogers for samples collected on the research cruise KOK 02-24 at Hawaii, L. Kerr for instruction and guidance on the SEM and confocal microscope, P. Schloss for support with the DOTUR and SONS programs, S. Simmons for support and instruction on qPCR, and E. Leadbetter for advice on the manuscript. This research was supported by a Ridge2K grant awarded to K.J.E. and W.B., in part by a NAI CAN awarded to M.L.S. and K.J.E., and also in part by a Project Development Award from WWU's Office of Research and Sponsored Programs to C.L.M.

Author Contributions Phylogenetic analyses were performed by C.M.S. for EPR samples, and B.N.O. and E.B. for Hawaii samples; microscopy was done by C.M.S. Biomass calculations were done by W.B. C.L.M. helped C.M.S. with qPCR analyses. M.L.S. supported Hawaii studies and the pipeline for phylogenetic studies. H.S. provided ship access for Hawaii studies. K.J.E. and W.B. developed and guided this project, and K.J.E., C.M.S. and B.N.O. developed and wrote the paper with input from co-authors.

Author Information Sequences are deposited in the GenBank database under accession numbers EU491521–EU491952 (for EPR basalts) and EU491020–EU491491 (for Hawaii basalts). Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to K.J.E. (kje@usc.edu).

METHODS

Sample collection. Basaltic seafloor lavas were collected from the EPR at 9° N (Supplementary Fig. 1a, b) during R/V *Atlantis* research cruises AT11-7 and AT11-20 and from around Hawaii (Supplementary Fig. 1a, c) during the *Ka'imikai-o-Kanaloa* cruise KOK 02-24 using the submersibles *Alvin* and *Pisces V*, respectively. Lavas were placed in bioboxes containing sterile fresh water, allowed to fill with ambient bottom sea water, sealed to minimize contamination, and brought to the surface for processing. Ambient EPR deep sea water was collected at the sea floor with Niskins bottles. Five basaltic lavas from the EPR and four Hawaii basalts (Supplementary Table 1) were used for analyses. **DNA extraction and sequencing.** The outer ≤ 1.5 cm glass portion of each was crushed to millimetre-sized or smaller fractions using aseptic techniques after collection. For the EPR basalts, subsamples were either processed at once or frozen (-80°C) for less than 24 h until DNA was extracted. Hawaii rock fragments were stored for ~ 3 yr at -80°C before DNA extraction. Environmental DNA was extracted from ~ 1 g crushed basalt using the Ultraclean soil DNA kit (MoBio Laboratories) after a modified manufacturer's protocol. Before bead beating, samples were incubated for 10 min at 70°C , after which $200\text{ }\mu\text{g}$ of polyadenylic acid (polyA) was added³¹. For the deep-seawater sample, 50 ml of fluid was filtered (0.2- μm -pore-size), and extractions were performed directly on the filter. The 16S rRNA region of environmental DNA was amplified using primer set 8F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTACCTGTTACGACTT-3').

Amplification products were purified using the QIAquick nucleotide removal kit (Qiagen). Clone libraries were constructed from amplification products as described in a previous publication²². Plasmid extractions (alkaline lysis) and sequencing (ABI v 3.1 BigDye terminator, Applied Biosystems) were performed at the Keck Facility of the JBPC, MBL, Woods Hole, Massachusetts.

Phylogenetic analysis. Sequence reads were edited and assembled using an automated pipeline, straw, from the JBPC (<http://jbpc.mbl.edu/computing-seqinformatics.html>). Finalized sequences were aligned using ARB (<http://www.arb-home.de/>). Chimaeric sequences were identified with Bellerophon³² (EPR basalts) and Mallard³³ (Hawaiian Island basalts). Putative chimaeras were checked against sequences of several closest relatives (determined by BLAST (Basic Local Alignment Search Tool))³⁴ using ARB, as well as with the RDP-II chimera detection program²³.

BLAST was used to for comparison against the GenBank 16S rRNA database for phylogenetic assignments. ARB was used to generate distance matrices using the neighbour-joining method, application of a filter for bacteria (positions 1218–42590), and the Jukes–Cantor correction. DOTUR²⁴ was used for rarefaction analysis and richness estimates using a 97% sequence similarity definition. Shared richness estimates were calculated using SONS²⁵. Sequences from the hydrothermal vent *in situ* growth chamber⁷ were accessed from GenBank. The farm soil sequences were provided with authors' permission⁹. Sargasso Sea data⁶, previously analysed in ref. 24, was downloaded from (<http://www.plantpath.wisc.edu/fac/joh/DOTUR/>).

Quantitative PCR. Real-time qPCR using the Taqman approach (Applied Biosystems) was used to quantify bacterial and archaeal 16S rDNA gene copies from environmental DNA. The qPCR assay to estimate Archaea used the primer set Arch349F (5'-GYGCASCAGKCGMGAAG-3') and Arch806R (5'-GGACT-ACVSGGGTATCTAAT-3') with probe Arch516F ((6-FAM)-5'-TGYCAGCC-GCCGCGGTAHACCVGC-3'-(Iowa Black FQ)), previously designed and tested in ref. 29. Primers and probe were synthesized by Integrated DNA Technologies. Primers 331F (5'-TCCTACGGGAGGCAGCAGT-3') and 797R (5'-GGACTACCAGGTATCTAATCTGT-3') and TaqMan probe (6-FAM)-

5'-CGTATTACCGCGGCTGCTGGCAG-3'-(TAMRA) (Thermo Electron Corporation) were used for bacterial DNA quantification²⁸.

Assays on an ABI Prism 7000 Sequence Detection System in 96-well optical grade plates used the following reagents: 1 \times TaqMan Universal PCR Master Mix (Applied Biosystems), 1 unit Platinum Taq (Invitrogen), 1 \times ROX (Invitrogen), 800 nM forward primer, 800 nM reverse primer, 200 nM Taqman probe, template (3 μl), and molecular grade water (total volume of 30 μl). Amplification conditions were 50°C for 2 min, 95°C for 10 min, and 45 cycles of 95°C for 15 s and 59°C for 3 min. All reactions (standards, samples, and blanks) were performed in triplicate and included a no template control.

Standard curves for quantifying gene copies were determined using purified PCR products from a mixture of archaeal and bacterial plasmid DNA of known concentration ($10\text{ ng }\mu\text{l}^{-1}$ each) isolated from microbial mats at hydrothermal vents along the Mariana Island Arc³⁵. For DNA quantification, serial dilutions ($1 \times 10^{-6} \times$ concentration) of PCR products were used to construct standard curves for Bacteria ($E = 94\%$, $r^2 = 0.99$, detection limit = 1.5×10^{-5} ng DNA per reaction) and Archaea ($E = 76\%$, $r^2 = 0.99$, detection limit = 1.5×10^{-5} ng DNA reaction⁻¹). DNA concentrations were converted to gene copy number assuming one double-stranded DNA molecule has a mass of 660 g mol^{-1} . Conversion factors of 4.1 (16S rDNA gene copies cell⁻¹) and 1.6 were used to convert bacterial and archaeal gene copies to cell numbers³⁰.

CARD-FISH and microscopy. Lavas were fixed in 4% (w/v) paraformaldehyde, rinsed, and stored in PBS:ethanol (1:1) at -20°C . Hybridizations were performed as described previously^{36,37}. Hybridizations were performed at 35°C with 55% formamide concentration. Bacteria or Archaea were targeted using horse-radish peroxidase-linked oligonucleotide probes EUB338-I, EUB338-II and EUB338-III (ref. 26) or ARCH915 (ref. 27), with Cy3-tyramides synthesized by Biomerns.net; SYBR Green I DNA stain (Invitrogen) was used as counterstain. An LSM510-META laser scanning confocal microscope (Carl Zeiss) based on an Axioskop 2FS fixed-stage upright with an Achromplan IR $\times 63/0.90\text{W}$ (water immersion) objective was used with a HeNe laser (543 nm wavelength) and Ar ion laser (488 nm wavelength) at the MBL Central Microscopy Center. LSM Image Browser v3.5 (Zeiss) was used to create three-dimensional image projections from z-stacks, and Adobe Photoshop was used for minimal image processing (cropping and adjusting contrast and/or brightness).

SEM samples were obtained at the MBL Microscopy Center with a JEOL 840 SEM operated at 15 kV.

31. Webster, G., Newberry, C. J., Fry, J. C. & Weightman, A. J. Assessment of bacterial community structure in the deep sub-seafloor biosphere by 16S rDNA-based techniques: a cautionary tale. *J. Microbiol. Methods* **55**, 155–164 (2003).
32. Huber, T., Faulkner, G. & Hugenholtz, P. Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* **20**, 2317–2319 (2004).
33. Ashelford, K. E., Chuzhanova, N. A., Fry, J. C., Jones, A. J. & Weightman, A. J. New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. *Appl. Environ. Microbiol.* **72**, 5734–5741 (2006).
34. Altschul, S. F. et al. Gapped BLAST and PSI-blast: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
35. Davis, R. E. & Moyer, C. L. Extreme spatial variability in microbial mat communities from submarine hydrothermal vents located at multiple volcanoes along the Mariana Island. *Eos. Trans. AGU* **86**, Fall Meet. Suppl. Abstract V51C-1509 (2005).
36. Pernthaler, A., Pernthaler, J. & Amann, R. Fluorescence in situ hybridization and catalyzed reporter deposition for the identification of marine bacteria. *Appl. Environ. Microbiol.* **68**, 3094–3101 (2002).
37. Sekar, R. et al. An improved protocol for quantification of freshwater *Actinobacteria* by fluorescence in situ hybridization. *Appl. Environ. Microbiol.* **69**, 2928–2935 (2003).

Functional genomic screen reveals genes involved in lipid-droplet formation and utilization

Yi Guo^{1,4*}, Tobias C. Walther^{1,5*}, Meghana Rao⁴, Nico Stuurman², Gohta Goshima^{2,†}, Koji Terayama⁴, Jinny S. Wong⁴, Ronald D. Vale^{2,6}, Peter Walter^{1,6} & Robert V. Farese Jr^{1,3,4}

Eukaryotic cells store neutral lipids in cytoplasmic lipid droplets^{1,2} enclosed in a monolayer of phospholipids and associated proteins^{3,4}. These dynamic organelles⁵ serve as the principal reservoirs for storing cellular energy and for the building blocks for membrane lipids. Excessive lipid accumulation in cells is a central feature of obesity, diabetes and atherosclerosis, yet remarkably little is known about lipid-droplet cell biology. Here we show, by means of a genome-wide RNA interference (RNAi) screen in *Drosophila* S2 cells that about 1.5% of all genes function in lipid-droplet formation and regulation. The phenotypes of the gene knock-downs sorted into five distinct phenotypic classes. Genes encoding enzymes of phospholipid biosynthesis proved to be determinants of lipid-droplet size and number, suggesting that the phospholipid composition of the monolayer profoundly affects droplet morphology and lipid utilization. A subset of the Arf1–COPI vesicular transport proteins also regulated droplet morphology and lipid utilization, thereby identifying a previously unrecognized function for this machinery. These phenotypes are conserved in mammalian cells, suggesting that insights from these studies are likely to be central to our understanding of human diseases involving excessive lipid storage.

We studied lipid-droplet formation in *Drosophila* Schneider 2 (S2) cells, a proven system for functional genomic studies with efficient gene inactivation by RNAi⁶. We induced lipid-droplet formation by incubation with 1 mM oleate for 24 h. Staining with 4,4-difluoro-1,3,5,7,8-pentamethyl-4-bora-3a,4a-diaza-s-indacene (BODIPY493/503) showed that droplet size, number and overall volume were increased (Fig. 1a); cellular triacylglycerol content increased sevenfold (Fig. 1b). We confirmed that BODIPY-stained fluorescent signals corresponded to lipid droplets with a red fluorescent protein mCherry⁷ fused with lipid storage droplet-1 (LSD1), which localizes exclusively to the surface of lipid droplets³ (not shown).

Imaging this process by time-lapse microscopy of BODIPY-labelled cells after oleate addition (Supplementary Movie 1) showed that droplet formation occurred in steps (Fig. 1c). First, increased numbers of small droplets formed in dispersed locations throughout the cell. Next, droplets increased in size and finally aggregated into one or several large clusters, resembling grapes. Electron microscopy confirmed the tight clustering of the droplets, which were often near the nucleus (Supplementary Fig. 3).

To unravel the molecular mechanisms governing this progression of changes during lipid-droplet formation, we performed a genome-wide RNAi screen in S2 cells (Fig. 2a). Images were acquired and examined by two independent observers, who scored them for alterations in droplet number, size and dispersion. The same data were

analysed computationally (Supplementary Methods). From visual screening, both observers identified 847 candidate genes with altered lipid-droplet morphology. To verify these genes and to minimize the misidentification of genes from off-target effects of RNAi treatments^{8,9}, RNAi experiments for these genes were repeated with a second, distinct set of double-stranded (ds)RNAs¹⁰. Visual analyses identified 132 genes whose knockdown consistently and repeatedly altered lipid-droplet morphology (Supplementary Table 1) and an additional 48 genes for which knockdown phenotypes were scored in two of three rounds (Supplementary Table 2). Computational analysis confirmed 86 of these 180 genes and added 47 genes with altered lipid-droplet morphology (Supplementary Table 3). Thus, we identified 227 genes (about 1.5% of the genome) that we conclude, with high confidence, affect lipid-droplet morphology. However, the high stringency of our criteria may have caused us to miss some genes involved in lipid-droplet morphology.

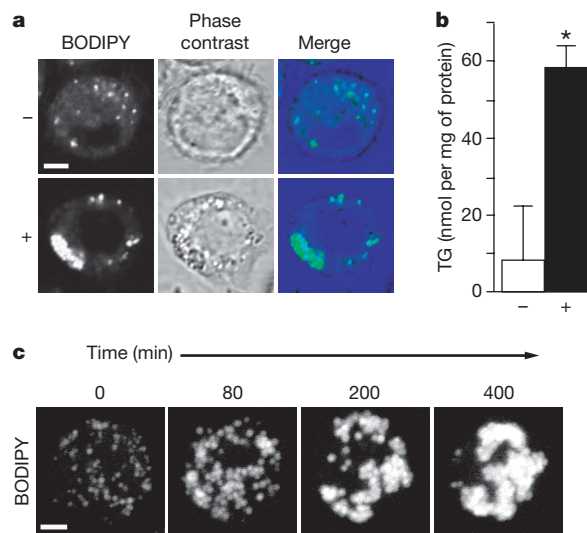


Figure 1 | Oleate increases the formation of lipid droplets in *Drosophila* S2 cells. **a**, S2 cells incubated for 24 h without (upper) or with (lower) 1 mM oleate. Staining with BODIPY, phase-contrast image, and overlay are shown. **b**, Oleate-loaded cells have an increased triacylglycerol (TG) content. Cells were incubated as in **a** and their TG contents were measured. Results are means and s.d. for three experiments; $P < 0.05$. **c**, Lipid-droplet formation occurs in steps. Cells treated as in **a** were stained with BODIPY. Single cells were followed by four-dimensional confocal time-lapse microscopy. Representative maximum projections of three-dimensional stacks at the indicated times are shown. Scale bars, 3 μ m.

¹Department of Biochemistry and Biophysics, ²Department of Cellular and Molecular Pharmacology, and ³Department of Medicine, University of California, San Francisco, California 94158, USA. ⁴Gladstone Institute of Cardiovascular Disease, San Francisco, California 94158, USA. ⁵Max-Planck Institute of Biochemistry, D-12852 Martinsried, Germany. ⁶Howard Hughes Medical Institute, University of California, San Francisco, California 94158-2517, USA. [†]Present address: Institute for Advanced Research, Nagoya University, 464-8601 Nagoya, Japan.

*These authors contributed equally to this work.

The 132 genes with striking phenotypes were categorized into five distinct phenotypic classes (Fig. 2b and Supplementary Table 1), which were validated for selected knockdowns by electron microscopy (Supplementary Fig. 3). Class I genes showed reduced numbers of droplets and included *midway* (encoding a diacylglycerol acyltransferase), subunits of the proteasome and the spliceosome, and several uncharacterized open reading frames. Class II genes gave smaller, more dispersed, droplets and included subunits of the COP9–signalosome complex, dynein, and RNA polymerase II subunits. Class III genes showed more dispersed droplets of slightly larger size and were members of the Arf1–COPI vesicular transport machinery. Class IV genes yielded highly condensed clusters of droplets and included members of the translational machinery. Class V genes contained one or a few very large droplets and included an orthologue of sterol regulatory element binding-protein (SREBP), a master transcriptional regulator of lipid metabolism, and SREBP cleavage activating protein (SCAP)¹¹. In *Drosophila*, the SREBP pathway is sensitive to and regulates phospholipid biosynthesis¹². This class also included *Cct1* and *Cct2*, which encode isoforms of phosphocholine cytidyltransferase, the enzyme that catalyses the rate-limiting step in phosphatidylcholine synthesis¹³, and CG2201, which is predicted to have choline kinase activity that phosphorylates and activates choline¹⁴. Thus, most class V genes were linked directly or indirectly to phospholipid biosynthesis.

To further explore how phospholipid metabolism regulates lipid-droplet formation, we characterized the *Cct1* and *Cct2* knockdowns. Larger droplets in *Cct* knockdowns could arise from a failure to form new droplets, forcing newly synthesized neutral lipids into a few large droplets, or from the fusion of independently formed droplets. To distinguish between these possibilities, we observed the dynamics of lipid-droplet formation by time-lapse microscopy (Supplementary Movie 2) and found evidence that the droplets fuse (Fig. 3a and Supplementary Fig. 2).

We then examined where CCT proteins act. In untreated cells, mCherry-tagged *Cct1* localized exclusively to the nucleus (Fig. 3b), similar to mammalian cytidyltransferase- α (CT- α)^{13,15}. After treatment with oleate, a significant portion of *Cct1* localized to the lipid-droplet surface. By contrast, similarly tagged *Cct2* localized to the

cytoplasm but was also concentrated on droplet surfaces after treatment with oleate. This marked translocation of CCT enzymes to the droplet surface may serve to provide adequate phosphatidylcholine to the phospholipid monolayers of growing lipid droplets. If so, the ratio of surface phospholipids to core neutral lipids may regulate lipid-droplet morphology: when phospholipids are limiting (as in *Cct1* or *Cct2* knockdowns), fusion is induced to decrease the surface-to-volume ratio of droplets. In fact, the content

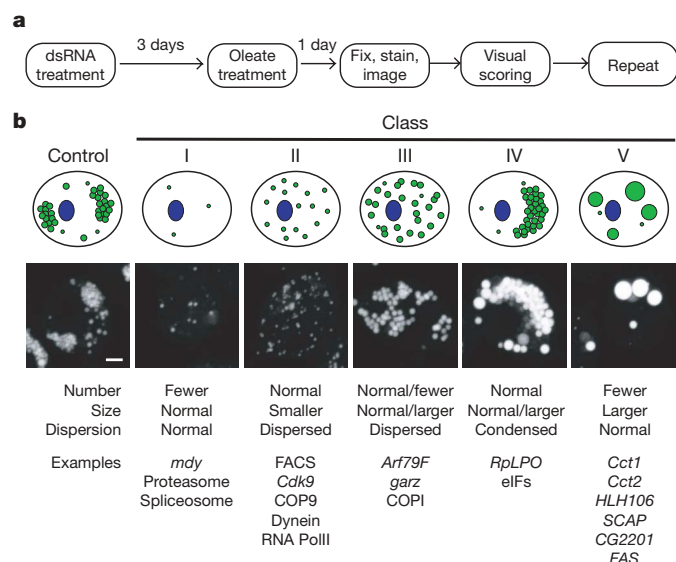


Figure 2 | Genome-wide screen identified genes regulating the formation of lipid droplets. **a**, Outline for strategy to screen for genes involved in lipid-droplet biogenesis. See the text for details. **b**, Genes involved in the screen for lipid-droplet biogenesis fall into distinct phenotypic classes. The 132 most striking phenotypes were classified according to lipid-droplet number, size and dispersion. From this classification, five major classes emerged; a graphic representation (top), an example image (middle) and some gene examples (bottom) are shown. Scale bar, 3 μ m.

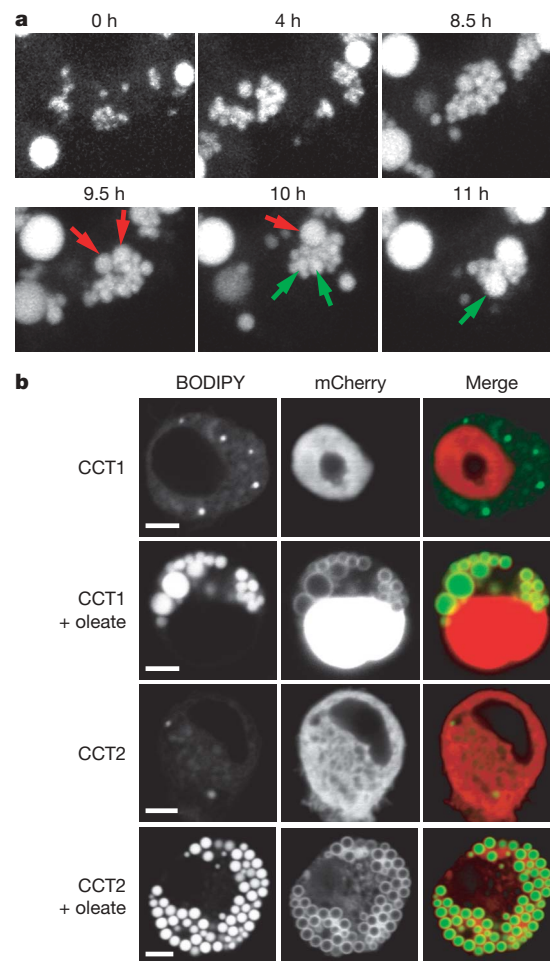


Figure 3 | Phosphatidylcholine content regulates the size and abundance of lipid droplets. **a**, Lipid-droplet formation induced by oleate in *Cct1* knockdown cells (Fig. 1a). Single cells were followed by time-lapse confocal microscopy. Representative projections revealed that droplets first proliferate normally (upper) and then fuse (lower). Examples of fusion are indicated (red and green arrows). **b**, CCT enzymes localize to the surface of droplets after induction with oleate. *Cct1* and *Cct2* were transiently expressed in S2 cells as amino-terminal mCherry-tagged fusion proteins and were stained and imaged before or after induction. BODIPY staining, mCherry fluorescence, and merge are shown. Scale bar, 3 μ m. **c**, *Cct1* knockdown cells have less phosphatidylcholine (PC) and more triacylglycerol (TG). S2 cells were treated with dsRNAs as indicated, loaded with oleate (as in Fig. 1a) and lysed. PC (upper) and TG (lower) levels in the extract were measured. Results are means and s.d. for three independent experiments. **, $P < 0.01$ versus control RNAi.

of phosphatidylcholine in cells with *Cct1* knockdown was decreased by about 60% (Fig. 3c, upper panel), and the triacylglycerol content was increased by about 40% (Fig. 3c, lower panel). The increase in triacylglycerol may reflect compensatory channelling of diacylglycerol into neutral lipids. A decreased phosphatidylcholine content would increase the relative amount of phosphatidylethanolamine in the droplet monolayer (as observed in flies lacking *Cct1* (ref. 16)), which itself may directly promote droplet fusion¹⁷. Our results suggest a model in which phosphatidylcholine availability is a crucial regulator of lipid-droplet size and number.

We also investigated class III genes, whose knockdowns showed slightly larger and more dispersed droplets. All class III genes were members of the Arf1–COPI machinery, including *Arf79F*, encoding an Arf1 family member, a gene encoding guanine nucleoside exchange factor (GEF), *garz*, and genes encoding components of the COPI coat. Similar effects were obtained by incubating cells with brefeldin A, a specific inhibitor of Arf1 exchange factors (Supplementary Fig. 5), and by expressing a dominant-negative version of *Arf79F*, encoding the T31N mutant, analogously to dominant-negative mutants for Ras or Ran¹⁸ (Fig. 4b). To test the specificity of this phenotype, we separately repeated RNAi knockdowns with

dsRNAs for *Drosophila* genes encoding six ARF proteins, three GEFs, two GTPase-activating proteins, and all COPI subunits. We also tested other coat proteins, such as clathrin subunits and components of the COPII coat (Fig. 4a). Only *Arf79F*, *garz* and six of eight members of the COPI coat (α -, β -, β' -, δ -, γ - and ζ -*Cop*) exhibited the class III phenotype, indicating that our screen identifies a highly specific subset of vesicular transport components. *Arf102F* knockdown gave a partial phenotype. This function of the Arf1–COPI machinery in lipid-droplet formation seems to have been evolutionarily conserved; similar phenotypes were found in yeast and human cells (Supplementary Fig. 4).

We next sought to determine whether Arf79F acts directly on lipid droplets. ARF proteins exchange rapidly between active (GTP-bound) and inactive (GDP-bound) forms, making it difficult to localize only the active form. However, Arf79F(T31N) binds its exchange factor tightly, and the distribution of the exchange factor is predicted to reflect the localization of active ARF protein. Expressed Arf79F(T31N) appeared diffusely in the cytosol but was enriched at the droplet surface (Fig. 4c). Thus, Arf79F may act at the lipid-droplet surface where, as for other ARF proteins, it interacts with its GEF (presumably encoded by *garz*) and recruits COPI

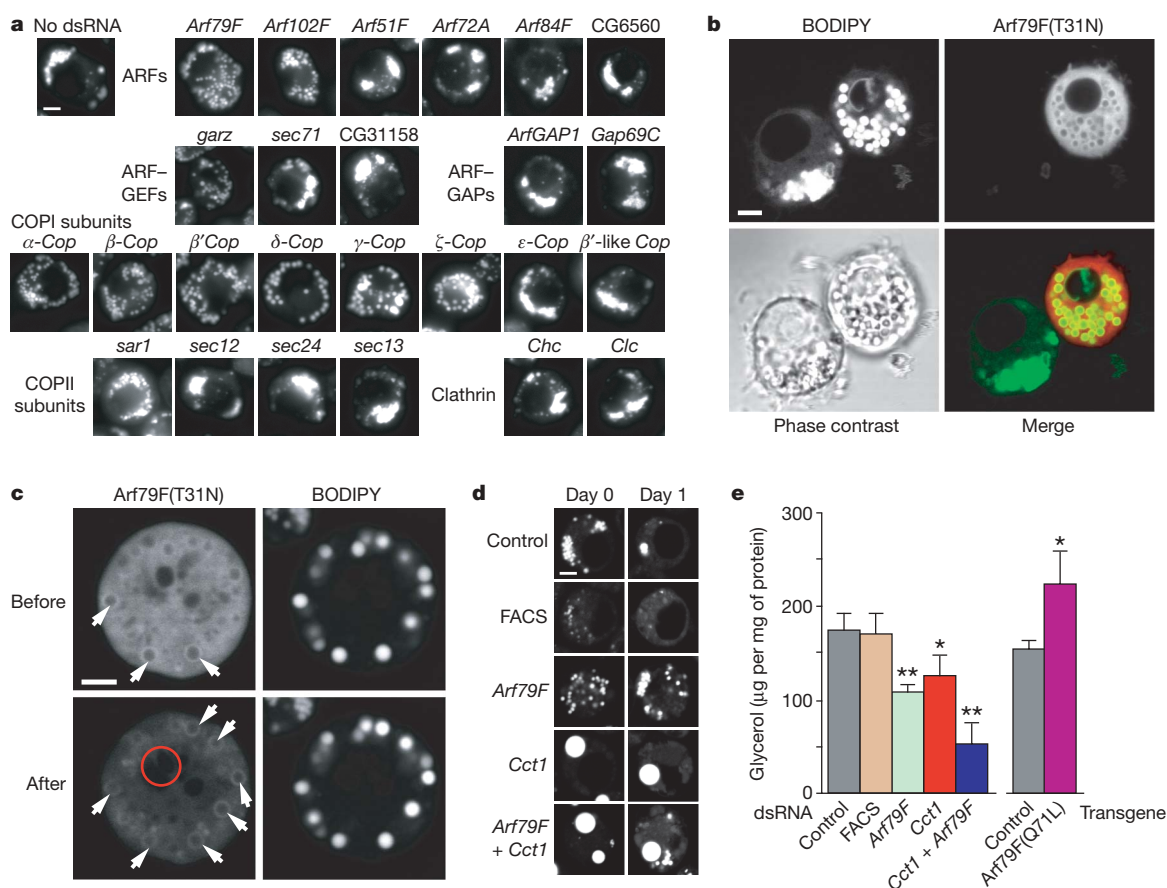


Figure 4 | Arf1–COPI complex members function in the formation of lipid droplets. **a**, Knockdowns of *Arf79F*, the GEF gene *gartenzweig* (*garz*) and specific subunits of the COPI coat affected lipid-droplet formation similarly. Representative images are shown. For descriptions of controls see Supplementary Information. GAP, GTPase-activating protein.

b, Arf79F(T31N) localizes to the droplet surface and causes a similar phenotype to *Arf79F* knockdown. Arf79F(T31N) expressed as a carboxy-terminal mCherry-tagged fusion protein in S2 cells was observed by confocal microscopy after loading with oleate and staining with BODIPY. A representative confocal midsection is shown for BODIPY (top left), mCherry fluorescence (top right) and a merge (bottom right).

c, Arf79F(T31N)–mCherry localizes to the droplet surface. The photobleached region (6 min) is indicated by a red circle. Arrows indicate the

association of Arf79F(T31N) with the surface of droplets. **d**, *Arf79F*, *Cct1* and double knockdowns lead to decreased lipolysis. S2 cells were treated with dsRNAs for three days as indicated, loaded with 1 mM oleate for one day, and imaged by confocal microscopy after staining with BODIPY (day 0, left panels). Representative confocal midsections are shown. Oleate was removed from the medium and the cells were starved for one day in serum-free medium to induce lipolysis (day 1, right). Scale bars, 3 μ m. FACS, gene encoding a long-chain-fatty-acid-CoA ligase (CG8732). **e**, *Arf79F*, *Cct1* and double knockdowns lead to decreased glycerol release to the medium. A transgene encoding Arf79F(Q71L) leads to increased release of glycerol. Experiments were as in **d**, and the glycerol released was measured. Results are means and s.d. for three independent experiments. *, $P < 0.05$, ** $P < 0.01$ versus control RNAi (left) and versus control transgene (right).

components. A recent *in vitro* study showed that Arf1 and several subunits of the COPI complex are recruited from the cytosol to purified lipid droplets in the presence of GTP- γ S (ref. 4). Although the recruitment of Arf1 to lipid droplets was reported to activate phospholipase D (ref. 19), we found no effect of phospholipase D knockdown on lipid-droplet formation in *Drosophila* cells (not shown).

The well-established functions of class III genes in Arf1–COPI-mediated vesicular transport^{20,21} implicate this machinery in a similar budding mechanism at the surface of lipid droplets, possibly to promote the budding-off of droplets during lipid mobilization. Lipolysis is associated with the break-up of larger droplets into smaller ones, presumably to provide more surface area for lipases²². We examined the effect of the *Arf79F* knockdown on lipolysis by inducing lipid-droplet formation (Fig. 4d) and then inducing lipid mobilization by incubation with serum-free medium lacking oleate. After 24 h, control cells had few droplets. By contrast, many droplets remained when *Arf79F* was inactivated. In addition, much less glycerol, a product of lipolysis, was released by cells lacking the Arf1–COPI machinery (Fig. 4e). Supporting a function of the Arf1–COPI machinery in lipolysis, more glycerol was released by cells expressing a dominant-active form of *Arf79F* encoding Arf79F(Q71L). These data indicate that the Arf1–COPI machinery is required for efficient lipolysis. Our data agree with a report showing that lipolysis in murine adipocytes is accompanied by a brefeldin A-sensitive process that is required for the mobilization of cholesterol from storage pools in droplets²³.

Increased droplet surface area during lipolysis would require more phospholipids in the surrounding monolayer. Because *Cct1* knockdown limits phosphatidylcholine amounts, we tested its effect on lipolysis. As predicted, *Cct1* knockdown markedly decreased the efficiency of lipolysis, as seen by lipid-droplet staining (Fig. 4d) and glycerol release (Fig. 4e). The effects of knockdowns of *Arf79F* and *Cct1* on lipolysis were additive, suggesting that these genes function independently.

Arf1–COPI complexes mediate retrograde vesicular trafficking of membranes and proteins from the Golgi apparatus to the endoplasmic reticulum²¹ and are also involved in vesicular transport processes from the *trans*-Golgi network and endosomes²¹. Notably, the members of the Arf1–COPI complex we identified were recently found in a *Drosophila* screen for genes involved in protein secretion and Golgi organization²⁴. Although the primary defect in class III knockdowns is as yet unknown, the phenotype on lipid-droplet formation is not likely to be an indirect consequence of inhibition of protein secretion. The effects are highly specific and are not observed with knockdowns of other proteins mediating secretory transport (endoplasmic reticulum translocation, COPII and clathrin). In addition, Arf79F is recruited to the lipid-droplet surface, as shown by us and others⁴, where it is presumably activated by the loading of GTP on its exchange factor.

Our study provides an initial systematic examination of the *Drosophila* genome to identify genes involved in lipid-droplet formation and utilization. Many genes that we identified sort to distinct classes of morphological changes, with each class containing functionally related proteins. These classes potentially link diverse processes, such as protein synthesis and degradation, the cell cycle, and organelle movement, with lipid-droplet biology. The variety of genes identified lends support to the emerging view of lipid droplets as dynamic organelles that are functionally connected to a variety of organelles and cellular processes, including the replication of intracellular pathogens such as *Chlamydia trachomatis* and hepatitis C^{25,26}. Many components of these processes are likely to be highly conserved across species. Our studies in S2 cells may therefore be directly relevant to cellular lipid storage in general, holding the promise of identifying pathways and mechanisms central to human diseases involving excessive lipid storage and to the engineering of cellular

lipid storage in organisms for the improved production of oils and biofuels.

METHODS SUMMARY

RNAi-mediated genomic screen. RNAi screening with University of California, San Francisco (UCSF), DmRNAi libraries versions 1 and 2 (15,683 genes) was as described¹⁰. Examination of selected genes demonstrated 80–90% mRNA knockdown (Supplementary Fig. 1); previous studies showed protein levels decreased by 80% (refs 27, 28). S2 cells (2.7×10^6 cells ml⁻¹) were treated with 1 μ g of dsRNA in 96-well plates for three days, and 1 mM oleic acid was added for 24 h. Cells were fixed for 1 h in 4% paraformaldehyde/PBS and stained for 1 h with BODIPY493/503 in PBS. Images were obtained with IC100 (Beckman) or ImageXpress Micro (Molecular Devices) automated microscopes and a 40 \times , 0.95 numerical aperture PlanApo dry objective lens (Nikon). Six image fields (400–600 cells per well) were acquired per RNAi experiment.

The primary visual screen was validated by treating S2 cells with an independent set of RNAs^{8–10}. Details of the visual and computational screening procedures are given in Supplementary Information. For all gene knockdowns, the penetrance of the phenotypes was high, affecting more than 70–80% of the cells.

Primers used for the original screen can be found at <http://rna.ucsf.edu/dropletscreen>, and primers for secondary validation can be found at http://mpibcms.biochem.mpg.de/en/rg/lipidrophe/absatz_01.html.

Protein localization. mCherry was from R. Tsien⁷. mCherry–Cct1, mCherry–Cct2 and Arf79F–mCherry expression vectors (actin promoter) were made with the Gateway system (Invitrogen). Arf79F(T31N)–mCherry and Arf79F(Q71L)–mCherry were generated by QuickChange II mutagenesis (Stratagene).

Lipid measurements. Cells were cultured in six-well plates (16 μ g of dsRNAs per well) for three days; 1 mM oleate was then added for 24 h. Cells were lysed in 50 mM Tris–HCl pH 7.4, 0.25 M sucrose; samples (250 μ g of protein) were assayed for lipids. Phosphatidylcholine was quantified with a colorimetric method²⁹. Triacylglycerol content was measured from extracted lipids by thin-layer chromatography³⁰.

Lipolysis studies. Cells were cultured with 1 mM oleate for 24 h. To stimulate lipolysis, oleate was removed and cells were cultured in serum-free medium for 24 h. Aliquots were analysed for glycerol (Sigma–Aldrich glycerol colorimetric assay) and protein (Bio-Rad D_C protein assay) content.

Received 4 February; accepted 25 March 2008.

Published online 13 April 2008.

- Bartz, R. *et al.* Lipidomics reveals that adiposomes store ether lipids and mediate phospholipid traffic. *J. Lipid Res.* **48**, 837–847 (2007).
- Brown, D. A. Lipid droplets: proteins floating on a pool of fat. *Curr. Biol.* **11**, R446–R449 (2001).
- Miura, S. *et al.* Functional conservation for lipid storage droplet association among Perilipin, ADPR, and TIP47 (PAT)-related proteins in mammals, *Drosophila*, and *Dicystostelium*. *J. Biol. Chem.* **277**, 32253–32257 (2002).
- Bartz, R. *et al.* Dynamic activity of lipid droplets: protein phosphorylation and GTP-mediated protein translocation. *J. Proteome Res.* **6**, 3256–3265 (2007).
- Martin, S. & Parton, R. G. Lipid droplets: a unified view of a dynamic organelle. *Nature Rev. Mol. Cell Biol.* **7**, 373–378 (2006).
- Ulvila, J. *et al.* Double-stranded RNA is internalized by scavenger receptor-mediated endocytosis in *Drosophila* S2 cells. *J. Biol. Chem.* **281**, 14370–14375 (2006).
- Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nature Biotechnol.* **22**, 1567–1572 (2004).
- Ma, Y., Creanga, A., Lum, L. & Beachy, P. A. Prevalence of off-target effects in *Drosophila* RNA interference screens. *Nature* **443**, 359–363 (2006).
- Kulkarni, M. M. *et al.* Evidence of off-target effects associated with long dsRNAs in *Drosophila melanogaster* cell-based assays. *Nature Methods* **3**, 833–838 (2006).
- Goshima, G. *et al.* Genes required for mitotic spindle assembly in *Drosophila* S2 cells. *Science* **316**, 417–421 (2007).
- Brown, M. S. & Goldstein, J. L. A proteolytic pathway that controls the cholesterol content of membranes, cells, and blood. *Proc. Natl Acad. Sci. USA* **96**, 11041–11048 (1999).
- Dobrosotskaya, I. Y., Seegmiller, A. C., Brown, M. S., Goldstein, J. L. & Rawson, R. B. Regulation of SREBP processing and membrane lipid production by phospholipids in *Drosophila*. *Science* **296**, 879–883 (2002).
- Kent, C. Regulatory enzymes of phosphatidylcholine biosynthesis: a personal perspective. *Biochim. Biophys. Acta* **1733**, 53–66 (2005).
- Morrison, D. K., Murakami, M. S. & Cleghon, V. Protein kinases and phosphatases in the *Drosophila* genome. *J. Cell Biol.* **150**, 57–62 (2000).
- Cornell, R. B. & Northwood, I. C. Regulation of CTP:phosphocholine cytidyltransferase by amphitropism and relocalization. *Trends Biochem. Sci.* **25**, 441–447 (2000).

16. Weber, U., Eroglu, C. & Mlodzik, M. Phospholipid membrane composition affects EGF receptor and Notch signaling through effects on endocytosis during *Drosophila* development. *Dev. Cell* **5**, 559–570 (2003).
17. Hafez, I. M. & Cullis, P. R. Roles of lipid polymorphism in intracellular delivery. *Adv. Drug Deliv. Rev.* **47**, 139–148 (2001).
18. Dascher, C. & Balch, W. E. Dominant inhibitory mutants of ARF1 block endoplasmic reticulum to Golgi transport and trigger disassembly of the Golgi apparatus. *J. Biol. Chem.* **269**, 1437–1448 (1994).
19. Nakamura, N., Banno, Y. & Tamiya-Koizumi, K. Arf1-dependent PLD1 is localized to oleic acid-induced lipid droplets in NIH3T3 cells. *Biochem. Biophys. Res. Commun.* **335**, 117–123 (2005).
20. Spang, A., Matsuoka, K., Hamamoto, S., Schekman, R. & Orci, L. Coatamer, Arf1p, and nucleotide are required to bud coat protein complex I-coated vesicles from large synthetic liposomes. *Proc. Natl Acad. Sci. USA* **95**, 11199–11204 (1998).
21. D'Souza-Schorey, C. & Chavrier, P. ARF proteins: roles in membrane traffic and beyond. *Nature Rev. Mol. Cell Biol.* **7**, 347–358 (2006).
22. Marcinkiewicz, A., Gauthier, D., Garcia, A. & Brasaemle, D. L. The phosphorylation of serine 492 of perilipin directs lipid droplet fragmentation and dispersion. *J. Biol. Chem.* **281**, 11901–11909 (2006).
23. Verghese, P. B., Arrese, E. L. & Soulages, J. L. Stimulation of lipolysis enhances the rate of cholesterol efflux to HDL in adipocytes. *Mol. Cell. Biochem.* **302**, 241–248 (2007).
24. Bard, F. *et al.* Functional genomics reveals genes involved in protein secretion and Golgi organization. *Nature* **439**, 604–607 (2006).
25. Kumar, Y., Cocchiari, J. & Valdivia, R. H. The obligate intracellular pathogen *Chlamydia trachomatis* targets host lipid droplets. *Curr. Biol.* **16**, 1646–1651 (2006).
26. Miyanari, Y. *et al.* The lipid droplet is an important organelle for hepatitis C virus production. *Nature Cell Biol.* **9**, 1089–1097 (2007).
27. Goshima, G. & Vale, R. D. The roles of microtubule-based motor proteins in mitosis: comprehensive RNAi analysis in the *Drosophila* S2 cell line. *J. Cell Biol.* **162**, 1003–1016 (2003).
28. Rogers, S. L., Wiedemann, U., Stuurman, N. & Vale, R. D. Molecular requirements for actin-based lamella formation in *Drosophila* S2 cells. *J. Cell Biol.* **162**, 1079–1088 (2003).
29. Hojjati, M. R. & Jiang, X. C. Rapid, specific, and sensitive measurements of plasma sphingomyelin and phosphatidylcholine. *J. Lipid Res.* **47**, 673–676 (2006).
30. Monetti, M. *et al.* Dissociation of hepatic steatosis and insulin resistance in mice overexpressing DGAT in the liver. *Cell Metab.* **6**, 69–78 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank I. Vasenkova and R. De Breuil for help with dsRNA synthesis; K. Warner for help with yeast work; M. Schuldiner, E. Griffis, T. Fazio, E. Herker, S. Stymne, B. Panning and M. Ott for reagents; D. B. Jones and G. Howard for assistance with manuscript preparation; G. Schoenhofer for web access for the database; members of the Farese, Vale and Walter laboratories for discussions; and D. Srivastava and S. Yamanaka for critical reading of the manuscript. This work was supported by a Freedom to Discover Award from Bristol-Myers Squibb and National Institutes of Health grant R21 DK078254-01 (to R.F.), a David and Mary Phillips postdoctoral fellowship award (to Y.G.), the Human Frontier Science Program Organization (T.C.W.), the Howard Hughes Medical Institute (P.W. and R.D.V.) and the J. David Gladstone Institutes.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to R.F. (bfarese@gladstone.ucsf.edu) or T.W. (twalthers@biochem.mpg.de).

LETTERS

Angiogenesis selectively requires the p110 α isoform of PI3K to control endothelial cell migration

Mariona Graupera¹, Julie Guillermet-Guibert¹, Lazaros C. Foukas¹, Li-Kun Phng², Robert J. Cain³, Ashreena Salpekar¹, Wayne Pearce¹, Stephen Meek⁴, Jaime Millan³, Pedro R. Cutillas¹, Andrew J. H. Smith⁴, Anne J. Ridley³, Christiana Ruhrberg⁵, Holger Gerhardt² & Bart Vanhaesebroeck¹

Phosphoinositide 3-kinases (PI3Ks) signal downstream of multiple cell-surface receptor types. Class IA PI3K isoforms¹ couple to tyrosine kinases and consist of a p110 catalytic subunit (p110 α , p110 β or p110 δ), constitutively bound to one of five distinct p85 regulatory subunits. PI3Ks have been implicated in angiogenesis^{2–5}, but little is known about potential selectivity among the PI3K isoforms and their mechanism of action in endothelial cells during angiogenesis *in vivo*. Here we show that only p110 α activity is essential for vascular development. Ubiquitous or endothelial cell-specific inactivation of p110 α led to embryonic lethality at mid-gestation because of severe defects in angiogenic sprouting and vascular remodelling. p110 α exerts this critical endothelial cell-autonomous function by regulating endothelial cell migration through the small GTPase RhoA. p110 α activity is particularly high in endothelial cells and preferentially induced by tyrosine kinase ligands (such as vascular endothelial growth factor (VEGF)-A). In contrast, p110 β in endothelial cells signals downstream of G-protein-coupled receptor (GPCR) ligands such as SDF-1 α , whereas p110 δ is expressed at low level and contributes only minimally to PI3K activity in endothelial cells. These results provide the first *in vivo* evidence for p110-isoform selectivity in endothelial PI3K signalling during angiogenesis.

p110 α and p110 β are ubiquitously expressed, whereas p110 δ is enriched in leukocytes^{6,7} and present at low levels in most other cell types^{4,5,8}. Ubiquitous knockout of p110 α (encoded by the *Pik3ca* gene) causes overall embryonic growth retardation and angiogenesis defects at mid-gestation^{2,9}. It is unclear, however, whether this is due to the loss of p110 α or to the resulting p85 upregulation² which has dominant-negative effects on all p110 isoforms¹⁰. We devised two genetic approaches to study selectively the role of p110 α activity in angiogenesis *in vivo*. First, we replaced endogenous p110 α with a kinase-dead allele (p110 α ^{D933A}) by introduction of a germline mutation in the p110 α gene, converting the ATP-binding DFG motif to AFG, without altering expression of p110 α and p85 (ref. 11). Second, we inactivated p110 α in endothelial cells by *Cre/loxP*-mediated recombination of a floxed p110 α gene (p110 α ^{lox}; Supplementary Fig. 1).

Growth retardation in p110 α ^{D933A/D933A} embryos was observed from embryonic day (E) 9.5 onwards, with full arrest at E10.5 and no live embryos at E12.5 (Supplementary Fig. 2a, b). All mutant embryos at E9.5 and E10.5 showed regular heartbeat and blood flow in central vessels (Supplementary Fig. 2c, d), suggesting that a functional cardiovascular system required for haemodynamic

force-dependent vascular remodelling¹² was established. However, blood accumulation in head and trunk, and the lack of large vitelline vessels (Supplementary Fig. 2e–g), indicated primary angiogenic remodelling defects as the likely cause of embryonic lethality. PI3K-dependent phosphorylation of Akt was absent in p110 α ^{D933A/D933A} embryos (Supplementary Fig. 3a), with no deregulation of p85 expression (Supplementary Fig. 3b).

At E9.5, major blood vessels such as the dorsal aorta and cardinal veins were present in p110 α ^{D933A/D933A} embryos (Supplementary Fig. 4), confirming that the initial stages of vascular development occurred in the absence of functional p110 α . However, the mutant dorsal aorta was thinner and the vascular plexus more primitive (Fig. 1a–d). By E10.5, vessels in the head (Fig. 1e–f) and trunk (Fig. 1g–h) were poorly remodelled and enlarged, with a primitive perineural vascular plexus and disorganized intersomitic vessels.

To study the endothelial cell-autonomous function of p110 α during angiogenesis, p110 α ^{lox/lox} mice were crossed to mice expressing *Cre* under the endothelial *Tie2* promoter¹³. This resulted in effective deletion of the floxed p110 α allele in *Tie2Cre/p110 α ^{lox/lox}* embryos (Supplementary Fig. 5a) and loss of expression of full-length p110 α without altering the stoichiometry of PI3K complexes (assessed in mouse embryonic fibroblasts from *Rosa26CreERT2/p110 α ^{lox/lox}* mice; Supplementary Fig. 5b). *Tie2Cre/p110 α ^{lox/lox}* embryos died before E12.5 (Supplementary Fig. 6a). At E10.5, they were similar in size to littermate controls (Supplementary Fig. 6b), but developed severe vascular abnormalities similar to those of p110 α ^{D933A/D933A} embryos, with defective remodelling of the primary yolk sac plexus (Supplementary Fig. 6c) and defective vascular patterning in head and trunk (Fig. 1i–p). These data suggest that the vascular defects in p110 α ^{D933A/D933A} embryos occurred independently of the overall growth retardation. Thus, p110 α is required in a cell-autonomous manner in endothelial cells to promote developmental angiogenesis and vascular patterning, and therefore embryonic survival.

We next studied vascular development upon inactivation of p110 β or p110 δ . Given that p110 β knockout results in lethality at the blastocyst stage¹⁴, we crossed conditional p110 β ^{lox/lox} mice¹⁵ to *Tie2Cre* mice, resulting in reduced lipid kinase activity in p110 β immunoprecipitates (Supplementary Fig. 7a, b). *Tie2Cre/p110 β ^{lox/lox}* mice were viable and fertile (Supplementary Fig. 7c), without obvious vascular defects (Supplementary Fig. 8a–d). Likewise, mice with inactive p110 δ (p110 δ ^{D910A/D910A})¹⁶ (Supplementary Fig. 9) also showed normal vascular development (Supplementary Fig. 8e–h). p110 α thus plays a unique role among class IA PI3K isoforms in embryonic vascular morphogenesis and remodelling.

¹Centre for Cell Signalling, Institute of Cancer, Queen Mary, University of London, Charterhouse Square, London EC1M 6BQ, UK. ²Vascular Biology Laboratory, Cancer Research UK London Research Institute, 44 Lincoln's Inn Fields, London WC2A 3PX, UK. ³King's College London, Randall Division of Cell and Molecular Biophysics, New Hunt's House, Guy's Campus, London SE1 1UL, UK. ⁴Gene Targeting Laboratory, The Institute for Stem Cell Research, University of Edinburgh, West Mains Road, Edinburgh EH9 3JQ, UK. ⁵Department of Cell Biology, Institute of Ophthalmology, University College London, London EC1V 9EL, UK.

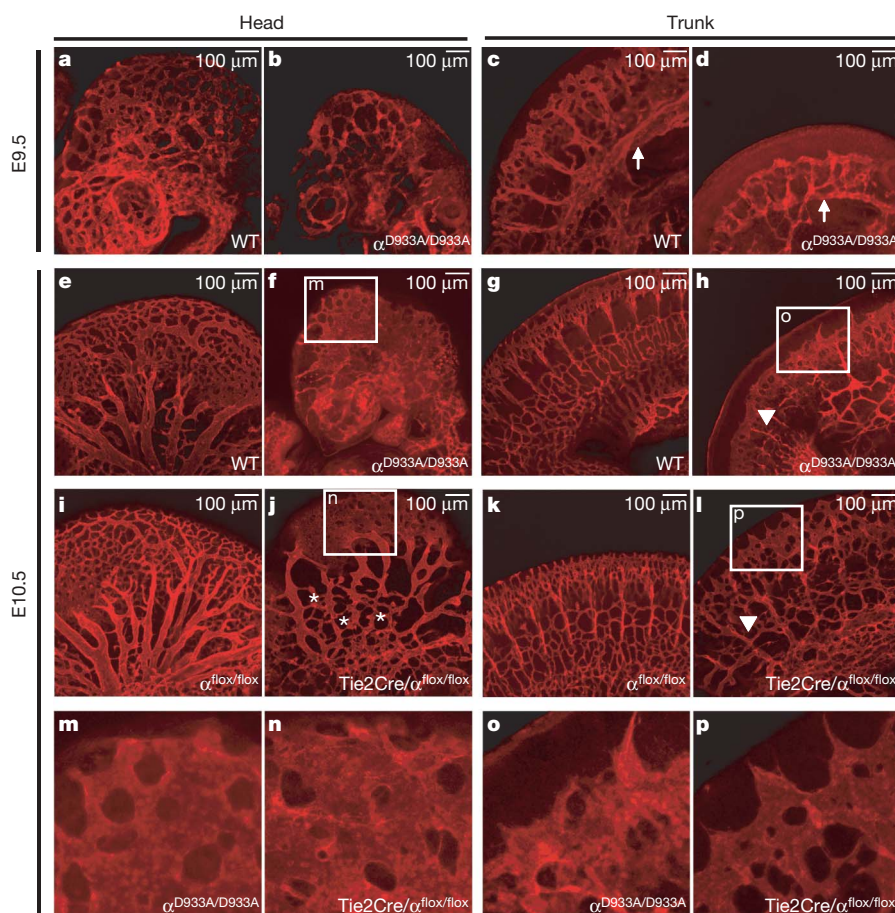
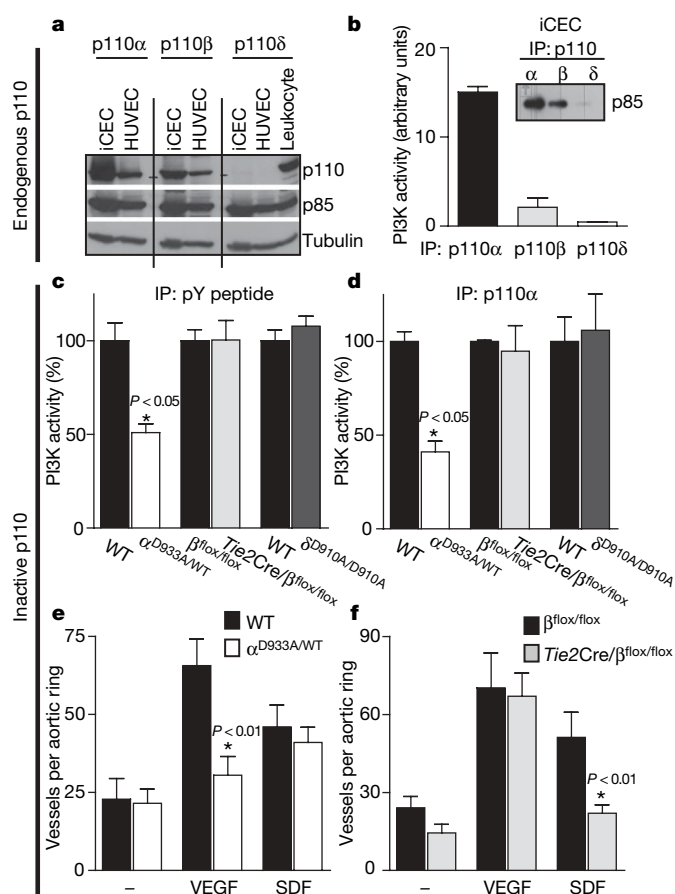


Figure 1 | Inactivation of p110 α in the germline (p110 $\alpha^{D933A/D933A}$) or in endothelial cells (Tie2Cre/p110 $\alpha^{flox/flox}$) leads to severe defects in angiogenic sprouting and vascular remodelling. Representative images of whole-mount E9.5 (a–d) and E10.5 (e–p) embryos labelled with endomucin (red), illustrating vascular plexus formation in the head and the trunk. At E9.5, p110 $\alpha^{D933A/D933A}$ embryos (b, d) showed a rudimentary and primitive vascular network with thinner dorsal aorta (white arrows). By E10.5, p110 $\alpha^{D933A/D933A}$ (f, h) and Tie2Cre/p110 $\alpha^{flox/flox}$ (j, l) exhibited abnormally enlarged blood vessels (white box), disorganized patterns (white triangle) and ectopic/irregular branching (asterisks). m–p, Higher magnification of the boxed areas shows highly similar features of immature plexus formation upon ubiquitous and conditional inactivation of p110 α in endothelial cells.



p110 α and p110 β were present in immortalized mouse cardiac endothelial cells (iCECs) and human umbilical vein endothelial cells (HUVECs) at levels similar to those of other cell types (Fig. 2a and data not shown), whereas p110 δ expression was very low, compared with leukocytes (Fig. 2a)^{4,5}. Among the class IA PI3K isoforms, p110 α was the most active in endothelial cells and in highly vascularized tissues (Fig. 2b and Supplementary Fig. 10). This is further supported by the observation that PI3K activity associated with a phosphoTyr peptide matrix (which binds all p85 species) or p110 α immunoprecipitates was reduced by approximately 50% in p110 $\alpha^{D933A/WT}$ lungs (Fig. 2c, d) and iCECs (Supplementary Fig. 11), but unaffected in Tie2Cre/p110 $\beta^{flox/flox}$ and p110 $\delta^{D910A/D910A}$ lungs (Fig. 2c). Conversely, the activity of p110 β and p110 δ was unaltered in p110 $\alpha^{D933A/WT}$ lungs (Supplementary Fig. 12). Thus, inhibition of one p110 isoform did not result in compensatory activity of the others.

We next studied the impact of inactivation of p110 α or p110 β on endothelial cell responses stimulated by various ligands. VEGF-A-stimulated microvessel outgrowth in aortic ring explants was reduced in p110 $\alpha^{D933A/WT}$ mice (Fig. 2e and Supplementary Fig. 13a), but not

Figure 2 | p110 α is the main provider of PI3K signalling in endothelial cells under basal and VEGF-A-stimulated conditions. a, PI3K subunit expression in wild-type iCECs and HUVECs. b, Lipid kinase activity and p85 content associated with p110 isoforms ($n \geq 3$ for each condition). IP, immunoprecipitate; pY, phosphoTyr. c, d, Lipid kinase activity associated with p85 (c) and p110 α (d) in adult lung ($n \geq 3$ for each genotype). Lung homogenates were absorbed onto an immobilized phosphoTyr peptide (which binds all p85 class IA PI3K regulatory subunits) or immunoprecipitated using p110 α antibody, followed by *in vitro* lipid kinase assay. e, f, Microvessel outgrowth of aortic rings from wild-type and p110 $\alpha^{D933A/WT}$ (e) and control and Tie2Cre/p110 $\beta^{flox/flox}$ (f) mice ($n \geq 6$ for each genotype and condition). Error bars, s.e.m.

in *Tie2Cre/p110 $\beta^{lox/lox}$* (Fig. 2f) and *p110 $\delta^{D910A/D910A}$* mice (Supplementary Fig. 14). *In vitro* tube formation was also impaired in *p110 $\alpha^{D933A/WT}$* iCECs (Supplementary Fig. 15a, b). These phenotypes could be mimicked in wild-type aortic rings and iCECs by LY294002, a pan-PI3K inhibitor, and by PI-103 (Supplementary Table 1), an inhibitor with selectivity for p110 α (Supplementary Figs 13c and 15c). Inhibitors of p110 β (TGX-155) or p110 δ (IC87114) had no significant effect on VEGF-A responses in these systems. VEGF-A-induced Akt phosphorylation was almost completely abrogated in primary CECs (pCECs; Supplementary Fig. 16) and lungs of *p110 $\alpha^{D933A/WT}$* mice (Supplementary Fig. 17).

A possible explanation for the minimal role of p110 β in VEGF-A signalling may relate to the recent observation that p110 β mainly couples to GPCRs¹⁵. *In vitro* microvessel outgrowth induced by the GPCR agonists SDF-1 α , IL8 or SIP was decreased in aortic rings from *Tie2Cre/p110 $\beta^{lox/lox}$* mice (Fig. 2f and Supplementary Fig. 18a), but not from *p110 $\alpha^{D933A/WT}$* mice (Fig. 2e). This was corroborated in wild-type aortic rings by pharmacological inhibition of p110 β (Supplementary Fig. 18b). Moreover, SDF-1 α - but not VEGF-A-induced phosphorylation of Akt was partially inhibited in *Tie2Cre/p110 $\beta^{lox/lox}$* pCECs (Supplementary Fig. 19).

PI3K activity is required for VEGF-A-dependent proliferation¹⁷, migration¹⁸ and survival^{19,20} of cultured endothelial cells. We therefore investigated whether p110 α inhibition affected any of these cellular processes. Partial inactivation of p110 α did not affect proliferation of iCECs and pCECs (Fig. 3a and Supplementary Fig. 20). *p110 $\alpha^{D933A/WT}$* iCECs also showed no defects in cell viability (Fig. 3b and Supplementary Fig. 21). However, clear alterations in migration were apparent (Supplementary Fig. 22). *p110 $\alpha^{D933A/WT}$* iCECs showed a 30% reduction in migration speed (12.2 ± 2 versus $7.6 \pm$

$1.2 \mu\text{m h}^{-1}$; $P < 0.05$) and total distance migrated (302.6 ± 95 versus $210 \pm 46.6 \mu\text{m}$). Likewise, migration of *p110 $\alpha^{D933A/WT}$* iCECs into scratch wounds in cell monolayers was impaired (Fig. 3c). These phenotypes were also observed in HUVECs in which p110 α expression was downregulated effectively (greater than 75%; Fig. 3d) and selectively (Fig. 3d) using short interfering RNA (siRNA). This led to reduced Akt phosphorylation (Fig. 3d) and a 40% decrease in migration speed (20.34 ± 11.88 versus $13.08 \pm 2.94 \mu\text{m h}^{-1}$; $P < 0.01$) and total distance migrated (663.8 ± 50 versus $374 \pm 32.4 \mu\text{m}$; $P < 0.01$), as well as a reduced monolayer scratch wound response (Fig. 3e), without altering proliferation (Fig. 3f) or survival (Fig. 3g).

We next asked whether reduced p110 α activity also affected endothelial cell migration *in vivo*. Vascular development in the postnatal retina²¹ and embryonic hindbrain²² are controlled by directed endothelial tip-cell migration towards VEGF-A. At postnatal day five (P5), *p110 $\alpha^{D933A/WT}$* retinas showed a significant delay in the outgrowth of the primary vascular plexus (Fig. 3h), indicating reduced migration of endothelial tip cells²³. Remodelling of arteries and veins, vessel diameter and branching density were unaffected, implying that spatial control of endothelial cell proliferation and survival was normal (Supplementary Fig. 23). Indeed, P5 wild-type and *p110 $\alpha^{D933A/WT}$* retinas showed similar endothelial 5-bromodeoxyuridine (BrdU) incorporation and lack of caspase-3 cleavage (Supplementary Fig. 24a, b). The presence of sprouts originating from hyaloid vessels only in the periphery of *p110 $\alpha^{D933A/WT}$* corroborated the delayed vascular development (Supplementary Fig. 24c). A similar delay in angiogenesis was observed in *p110 $\alpha^{D933A/WT}$* embryonic hindbrain: at E11.5, the density of vascular sprouts entering the pial side of the hindbrain and the density of the vascular

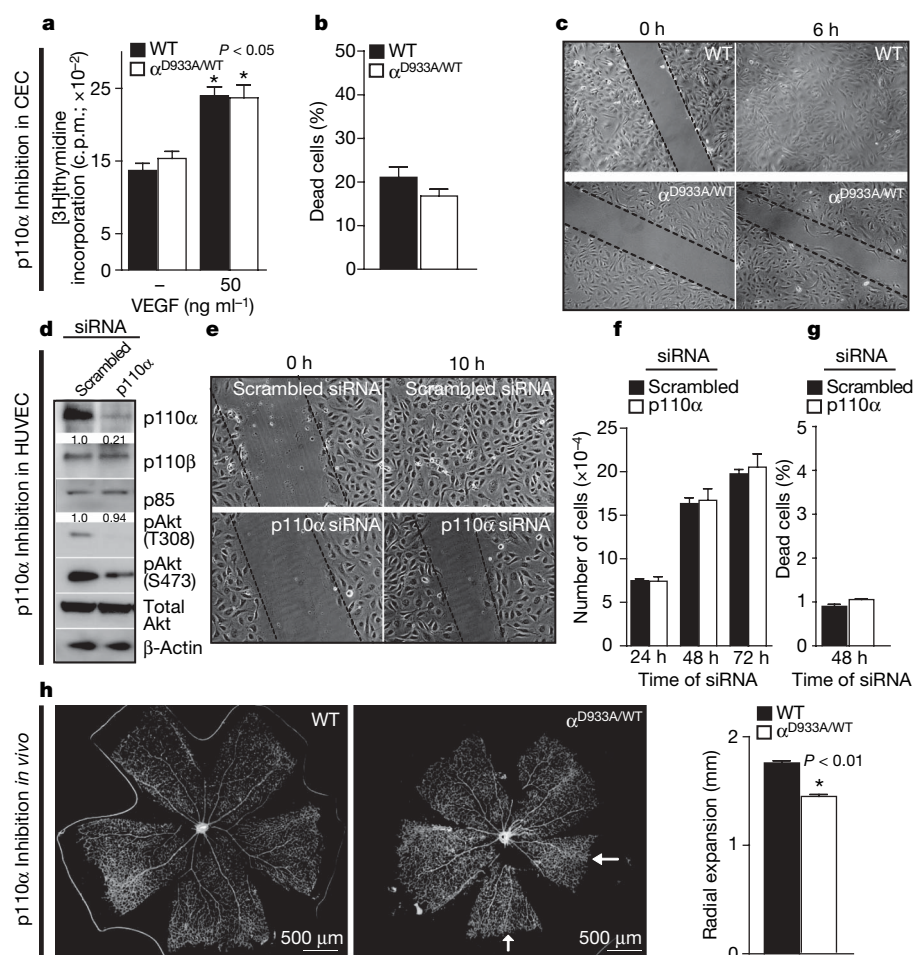


Figure 3 | p110 α controls endothelial cell migration *in vitro* and *in vivo*. **a**, VEGF-A-induced proliferation in iCECs. **b**, Cell viability of iCECs after 24 h starvation of serum, assessed by flow cytometry. **c**, Phase-contrast images of iCEC monolayers, immediately and 6 h after scratching. **d–g**, Effect of transfection of HUVECs with scrambled or p110 α siRNA on PI3K subunit expression and Akt phosphorylation (the numbers below represent densitometric evaluation of the data, expressed as a fold over β -actin) (**d**), monolayer scratch wound healing (**e**), cell numbers (**f**) and viability (**g**). Assays were performed 48 h (**d**, **g**) or 24 h (**e**) after transfection. At least three independent transfections were done for each siRNA (**d–g**). **h**, Whole-mount of P5 retinas stained for isolectin-B4 (with quantification of radial vessel expansion; $n \geq 4$ for each genotype). White arrows point towards reduced outgrowth of the primitive vascular plexus. Error bars, s.e.m.

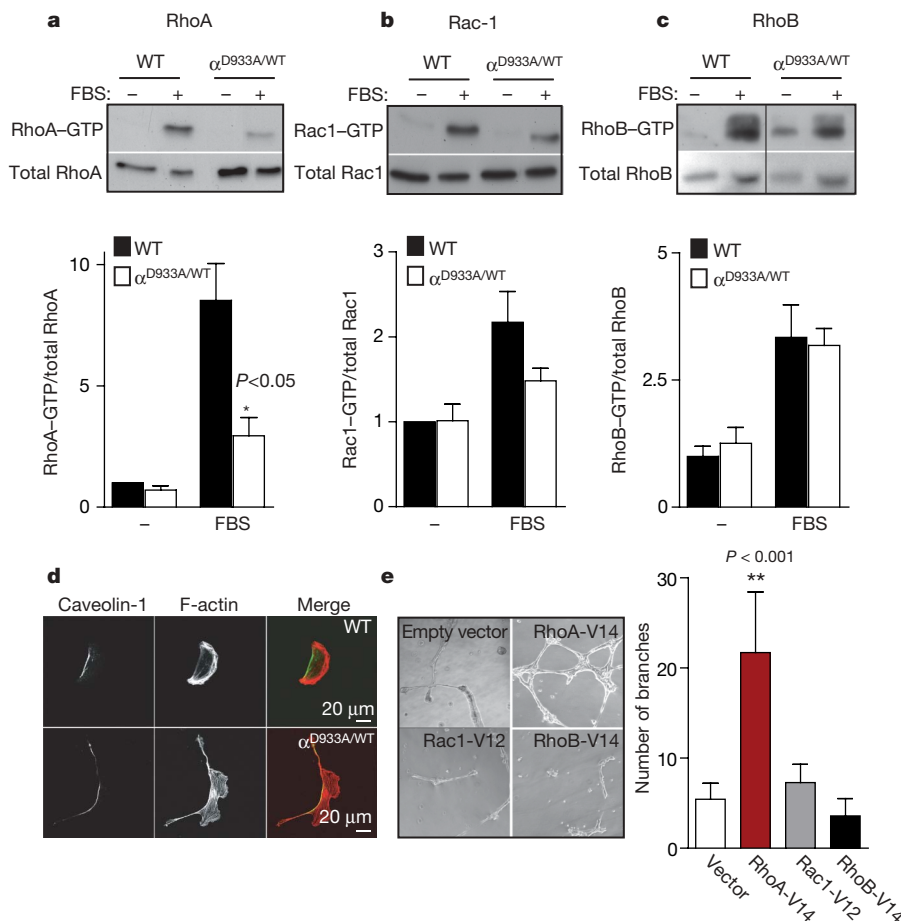


Figure 4 | p110 α is a positive regulator of RhoA in endothelial cells.

a–d, Activation of RhoA (**a**), Rac1 (**b**) and RhoB (**c**) in starved and serum-stimulated iCECs. Quantification of $n \geq 3$ experiments is shown.

d, Immunostaining of caveolin-1 (green) and F-actin (red) in iCECs.

e, Representative Matrigel micrograph of capillary network formation of

p110 $\alpha^{D933A/WT}$ iCECs transfected with constitutively active RhoA-V14, Rac1-V12 or RhoB-V14. The graph shows quantification of branch points in iCECs transfected with the indicated GTPases, seeded 24 h after transfection on Matrigel in the presence of VEGF ($n \geq 4$ for each condition). Error bars, s.e.m.

network in the subventricular zone were reduced (Supplementary Fig. 25a), with unaltered BrdU incorporation (Supplementary Fig. 25b). Thus, p110 α in endothelial cells appears primarily to regulate cell migration *in vitro* and *in vivo*.

We next monitored the activity of Rho GTPases, key players in cell migration. Genetic or pharmacological inactivation of p110 α reduced the activity of RhoA, both under basal (Supplementary Fig. 26) and serum-stimulated conditions (Fig. 4a and Supplementary Fig. 27), with minor effects on Rac-1 and RhoB (Fig. 4b, c). p110 $\alpha^{D933A/WT}$ iCECs produced long tails at the rear (Fig. 4d and Supplementary Fig. 28), suggesting a defect in cell detachment, known to be controlled by RhoA²⁴. Knockdown of p110 α expression in HUVECs also led to the generation of long tails (Supplementary Fig. 29). Expression in p110 $\alpha^{D933A/WT}$ iCECs of constitutively active RhoA-V14, but not Rac1-V12 or RhoB-V14, partially rescued the defect in tube formation (Fig. 4e). These data show that p110 α is a positive regulator of RhoA in endothelial cells and suggest that p110 α controls endothelial cell migration by regulating RhoA activity.

Our studies reveal a surprising degree of isoform selectivity among the class IA PI3K members in developmental angiogenesis *in vivo* and endothelial cell functions *in vitro*: p110 α as a single intermediate in VEGF-A signalling, p110 β as a key effector of GPCRs and a minimal contribution for the leukocyte-enriched PI3K isoform, p110 δ . The selective importance of p110 α may be partly explained by the prominent expression and activity of p110 α in endothelial cells. Despite the many functions of VEGF/VEGFR signalling in endothelial cells, inactivation of p110 α appears to impair selectively cell migration, primarily through RhoA-dependent signals required for

tail release during cell motility. The observed failure in vascular remodelling in the absence of defects in endothelial cell proliferation or survival supports the idea that endothelial cell migration may have a key role not only in angiogenic sprouting, but also in the subsequent remodelling²⁵.

METHODS SUMMARY

Four different mouse lines were used in this study (p110 $\alpha^{D933A/D933A}$, p110 $\alpha^{flox/flox}$, p110 $\beta^{flox/flox}$, p110 $\delta^{D910A/D910A}$). The conditional mouse lines were crossed to mice expressing Cre under the control of the endothelial *Tie2* promoter¹³. Embryos were obtained from timed pregnant females, with the day of the presence of copulation plug being considered as E0.5.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 5 November 2007; accepted 5 March 2008.

Published online 30 April 2008.

- Vanhaesebroeck, B., Leevers, S. J., Panayotou, G. & Waterfield, M. D. Phosphoinositide 3-kinases: a conserved family of signal transducers. *Trends Biochem. Sci.* **22**, 267–272 (1997).
- Bi, L. *et al.* Proliferative defect and embryonic lethality in mice homozygous for a deletion in the p110 α subunit of phosphoinositide 3-kinase. *J. Biol. Chem.* **274**, 10963–10968 (1999).
- Jiang, B. H., Zheng, J. Z., Aoki, M. & Vogt, P. K. Phosphatidylinositol 3-kinase signaling mediates angiogenesis and expression of vascular endothelial growth factor in endothelial cells. *Proc. Natl Acad. Sci. USA* **97**, 1749–1753 (2000).
- Geng, L. *et al.* A specific antagonist of the p110 δ catalytic component of phosphatidylinositol 3'-kinase, IC486068, enhances radiation-induced tumor vascular destruction. *Cancer Res.* **64**, 4893–4899 (2004).

5. Puri, K. D. *et al.* Mechanisms and implications of phosphoinositide 3-kinase δ in promoting neutrophil trafficking into inflamed tissue. *Blood* **103**, 3448–3456 (2004).
6. Vanhaesebroeck, B. *et al.* P110delta, a novel phosphoinositide 3-kinase in leukocytes. *Proc. Natl Acad. Sci. USA* **94**, 4330–4335 (1997).
7. Chantray, D. *et al.* p110delta, a novel phosphatidylinositol 3-kinase catalytic subunit that associates with p85 and is expressed predominantly in leukocytes. *J. Biol. Chem.* **272**, 19236–19241 (1997).
8. Sawyer, C. *et al.* Regulation of breast cancer cell chemotaxis by the phosphoinositide 3-kinase p110 δ . *Cancer Res.* **63**, 1667–1675 (2003).
9. Lelievre, E. *et al.* Deficiency in the p110alpha subunit of PI3K results in diminished Tie2 expression and Tie2(–/–)-like vascular defects in mice. *Blood* **105**, 3935–3938 (2005).
10. Ueki, K., Algenstaedt, P., Mauvais-Jarvis, F. & Kahn, C. R. Positive and negative regulation of phosphoinositide 3-kinase-dependent signaling pathways by three different gene products of the p85 α regulatory subunit. *Mol. Cell. Biol.* **20**, 8035–8046 (2000).
11. Foukas, L. C. *et al.* Critical role for the p110 α phosphoinositide-3-OH kinase in growth and metabolic regulation. *Nature* **441**, 366–370 (2006).
12. Lucitti, J. L. *et al.* Vascular remodeling of the mouse yolk sac requires hemodynamic force. *Development* **134**, 3317–3326 (2007).
13. Kisanuki, Y. Y. *et al.* Tie2-Cre transgenic mice: a new model for endothelial cell-lineage analysis *in vivo*. *Dev. Biol.* **230**, 230–242 (2001).
14. Bi, L., Okabe, I., Bernard, D. J. & Nussbaum, R. L. Early embryonic lethality in mice deficient in the p110 β catalytic subunit of PI 3-kinase. *Mamm. Genome* **13**, 169–172 (2002).
15. Guillermet-Guibert, J. *et al.* The p110 β isoform of phosphoinositide 3-kinase signals downstream of G protein-coupled receptors and is functionally redundant with p110 γ . *Proc. Natl Acad. Sci. USA* (in the press).
16. Okkenhaug, K. *et al.* Impaired B and T cell antigen receptor signaling in p110 δ PI 3-kinase mutant mice. *Science* **297**, 1031–1034 (2002).
17. Dayanir, V., Meyer, R. D., Lashkari, K. & Rahimi, N. Identification of tyrosine residues in vascular endothelial growth factor receptor-2/FLK-1 involved in activation of phosphatidylinositol 3-kinase and cell proliferation. *J. Biol. Chem.* **276**, 17686–17692 (2001).
18. Gille, H. *et al.* A repressor sequence in the juxtamembrane domain of Flt-1 (VEGFR-1) constitutively inhibits vascular endothelial growth factor-dependent phosphatidylinositol 3'-kinase activation and endothelial cell migration. *EMBO J.* **19**, 4064–4073 (2000).
19. Gerber, H. P., Dixit, V. & Ferrara, N. Vascular endothelial growth factor induces expression of the antiapoptotic proteins Bcl-2 and A1 in vascular endothelial cells. *J. Biol. Chem.* **273**, 13313–13316 (1998).
20. Adini, I. *et al.* RhoB controls Akt trafficking and stage-specific survival of endothelial cells during vascular development. *Genes Dev.* **17**, 2721–2732 (2003).
21. Gerhardt, H. *et al.* VEGF guides angiogenic sprouting utilizing endothelial tip cell filopodia. *J. Cell Biol.* **161**, 1163–1177 (2003).
22. Ruhrberg, C. *et al.* Spatially restricted patterning cues provided by heparin-binding VEGF-A control blood vessel branching morphogenesis. *Genes Dev.* **16**, 2684–2698 (2002).
23. Gerhardt, H. & Betsholtz, C. How do endothelial cells orientate? *EXS* **94**, 3–15 (2005).
24. Worthylake, R. A., Lemoine, S., Watson, J. M. & Burridge, K. RhoA is required for monocyte tail retraction during transendothelial migration. *J. Cell Biol.* **154**, 147–160 (2001).
25. Hughes, S. & Chan-Ling, T. Roles of endothelial cell migration and apoptosis in vascular remodeling during development of the central nervous system. *Microcirculation* **7**, 317–333 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank F. Ramadani and K. Okkenhaug (Babraham Institute, Cambridge), E. Cernuda (Hospital Universitario Central de Asturias), T. Makinen (Cancer Research UK London Research Institute), K. Hodivala-Dilke, A. Reynolds and G. D'Amico (Institute of Cancer, Queen Mary, University of London), P. Villalonga (Universitat de les Illes Balears, Spain) and members of the Vanhaesebroeck laboratory (especially N. Osborne, C. See and M. Whitehead) for help and advice, E. Wagner (Research Institute of Molecular Pathology, Vienna), E. Dejana (Institute of Molecular Oncology, Milan), G. Balconi (Mario Negri Institute for Pharmacological Research, Milan), M. Yanagisawa (University of Texas Southwestern Medical Center, Dallas), D. Vestweber (Max-Planck Institute, Muenster), C. Rommel, M. Camps and T. Ruckle (Merck-Serono, Geneva) and Piramed (Slough, UK) for mice and reagents. Personal support was from EMBO (M.G., J.G.-G.), Cancer Research UK (M.G.) and the Fondation pour la Recherche Médicale and the European Union Marie Curie (J.G.-G.). Work in the Vanhaesebroeck laboratory was supported by the Ludwig Institute for Cancer Research Institute, the Biotechnology and Biological Sciences Research Council (BB/C505659/1), the Association for International Cancer Research, European Union (FP6-502935), Cancer Research UK and Barts and the London Charity. R.J.C. is supported by an Association for International Cancer Research grant to A.J.R. (07-0173). L.-K.P. and H.G. are supported by Cancer Research UK.

Author Contributions All authors designed research and analysed data. M.G., J.G.-G., L.C.F., L.-K.P., R.J.C., A.S., W.P., S.M. and P.R.C. performed research. M.G., H.G. and B.V. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Correspondence and requests for materials should be addressed to H.G. (Holger.gerhardt@cancer.org.uk) or B.V. (bart.vanh@qmul.ac.uk).

METHODS

Reagents. VEGF165 (referred to as VEGF-A in the text), SDF-1 α , IL8 and S1P were from Preprotech. HUVECs and EBM-2 medium were obtained from Lonza. Catalogue numbers are shown between brackets. Antibodies to Akt/PKB (9272), pT308-Akt (4056), pS473-Akt (9271) and caspase-3 (9664) were from Cell Signalling Technologies. Antibodies to CD31 (550274), paxillin (610051) and CD105 (550546) were from BD Pharmingen. Antibodies to p110 α were either monoclonal clone U3A or rabbit anti-p110 α (made in house). Antibodies to p110 δ were also made in house. Anti-pan-p85 (06-195), anti-pTyr (4G10, catalogue number 16-101) and anti-Rac (05-389) were from Upstate. Antibodies to p110 β (sc-602), VEGFR-2 (sc-504), pErk1/2 (sc-7383), caveolin-1 (sc-894), RhoB (sc-180) and RhoA (sc-418) were from Santa Cruz Biotechnology. Anti-isolectin B4 (L-2140), anti-laminin, anti-tubulin (T5168) and anti- β -actin (A5441), phalloidin-TRITC (P1951) and all chemicals, unless otherwise stated, were from Sigma-Aldrich. BrdU, anti-BrdU Alexa Fluor 488, isolectin IB₄ Alexa Fluor 568 and secondary antibodies conjugated to Alexa-488, and Alexa-594 were from Molecular Probes. Anti-NG2 (AB5320) was from Chemicon International. Rat anti-endothelial mouse monoclonal antibody was provided by D. Vestweber. Secondary antibodies conjugated to fluorescein isothiocyanate (FITC), Cy5 and TRITC were from Jackson ImmunoResearch Labs.

Whole-mount embryo immunofluorescence. Freshly isolated embryos were fixed overnight in 4% paraformaldehyde at 4 °C, washed in PBT (PBS, 0.1% Tween 20) and subjected to dehydration in increasing methanol concentration (50%, 80% methanol/PBT, 100% methanol) followed by rehydration in decreasing methanol concentration (80%, 50% methanol/PBT, PBT). After washing in Pblec (PBS pH 6.8, 1% Tween 20, 1 mM CaCl₂, 1 mM MgCl₂, 0.1 mM MnCl₂), the embryos were incubated overnight at 4 °C in the presence of rat anti-mouse endomucin diluted 1:20 in Pblec. After five washes in PBT, embryos were incubated overnight at 4 °C with goat anti-rat Alexa 568 diluted 1:100 in PBS, 0.5% BSA and 0.25% Tween 20, followed by washing in PBT and post-fixation in 4% paraformaldehyde (PFA) before analysis. Confocal laser scanning microscopy was performed with an LSM 510 (Zeiss) mounted over an Axioplan microscope (Zeiss) using a $\times 10$ objective.

PI3K isoform expression and lipid kinase assay. PBS was perfused through the heart of terminally anaesthetized mice, followed by harvesting and freezing of the tissues in liquid nitrogen. PI3K isoform expression and lipid kinase assays were performed as previously described²⁶. Three independent samples for each genotype were used.

Aortic ring assay. Thoracic aortic rings (1 mm; $n \geq 6$ for each genotype and condition) were embedded on top of 50- μ l rat collagen R from rat tail (Serva) and overlaid with 100 μ l of OptiMEM (GIBCO), with or without VEGF-A (30 ng ml⁻¹), SDF-1 α (30 ng ml⁻¹), IL8 (30 ng ml⁻¹) or S1P (30 ng ml⁻¹). Microvessel outgrowth was visualized by time-lapse phase contrast microscopy, and the number of vessels growing from each aortic ring was counted on day 8. Representative images were taken using a KPM1E/K-S10 CCD camera (Hitachi Denshi). For PI3K inhibitor studies, aortic rings were incubated with inhibitors or vehicle for the whole duration of the study. Medium was changed every other day in the presence of fresh PI3K inhibitors. Quantification was performed as shown in Supplementary Fig. 13b.

CEC isolation. CECs were isolated from 8- to 12-week-old mice as described²⁷. Three independent cell lines were derived for each genotype. Briefly, hearts were aseptically removed, diced and digested with collagenase A (Roche). CECs were isolated by immunoselection using anti-CD105 (MJ7/18 monoclonal antibody)-conjugated beads, followed by plating of the cell suspension on 0.1% gelatin-coated dishes in DMEM (Invitrogen) supplemented with 10% fetal bovine serum (Helena Biosciences), 10 U ml⁻¹ heparin and 50 μ g ml⁻¹ endothelial cell growth factor. Two days later, the cells were infected for 4 h with retrovirus encoding polyoma middle T antigen. After overnight incubation in endothelial cell culture medium, a second 4 h round of infection with polyoma middle T antigen retrovirus was performed. Cultures were then washed, incubated for 2 weeks, and cells subjected to a second round of immuno-isolation as described above. Expression of various markers was assessed by PCR, or by fluorescence-activated cell sorting in the case of CD105. CECs of both genotypes were found to be positive for VEGFR1, VEGFR2, TIE1, TIE2, VE-cadherin, CD34 and CD31 (data not shown). Primary CEC isolation was as described above, with a single positive selection and assessment of CD105 expression by immunofluorescence.

Stimulation of pCECs with soluble ligands. For stimulation, early passage (P4) pCECs were seeded in 6 cm dishes (2.5×10^5 cells per dish), starved in DMEM without fetal bovine serum (FBS) and antibiotics for 24 h, followed by stimulation with VEGF (50 ng ml⁻¹ and 100 ng ml⁻¹), SDF-1 α (30 ng ml⁻¹) or FBS (10%) for 10 min, unless stated otherwise. Quantification of at least three independent experiments was performed.

In vitro endothelial tube formation assay. Growth-factor-reduced Matrigel matrix (BD Pharmingen, catalogue number 354230) was plated evenly in a 12-well plate (150 μ l per well) and incubated at 37 °C for 30 min before adding 75,000 CECs per well. Cells were overlaid with 2% serum medium with or without 50 ng ml⁻¹ VEGF-A. Sixteen hours after seeding, photographs of representative $\times 10$ fields were taken ($n \geq 5$ for each condition and genotype), and endothelial tubes were quantified by counting branches.

In vivo stimulation with VEGF-A. Saline or mouse VEGF-A (0.2 μ g g⁻¹) was injected into the tail vein of 8- to 12-week-old mice ($n \geq 6$ for each genotype and condition), followed 5 min later by harvesting and freezing of the tissues in liquid nitrogen, as described²⁸. Tissues were homogenized and re-suspended in different buffers according to the application. After removal of insoluble material by centrifugation at 12,000g for 10 min at 4 °C, 100 μ g cell lysate was analysed by western blot analysis, or 1 mg of lysate used to immunoprecipitate pTyr followed by western blot analysis. The Akt activity assay was performed as described²⁹.

Proliferation assay. Immortalized CECs and primary CECs were incubated at 10^5 and 2×10^5 cells, respectively, per well in 96-well plate, deprived of FBS for 16 h, followed by stimulation with or without 50 ng ml⁻¹ VEGF in medium (DMEM, 2% FBS, penicillin/streptomycin) containing [³H]thymidine. Twenty-four hours later, the cells were harvested and [³H]thymidine incorporated into DNA measured by scintillation counting.

Cell viability assays. Cells, grown at 5×10^5 cells per 10-cm dish for 24 h, were starved of FBS for 20 h, followed by analysis of cell viability by flow cytometry using the Annexin V-FITC apoptosis kit I (BD Pharmingen).

Caspase activity assay. Cells, grown at 3×10^5 cells per 10-cm dish for 24 h, were starved of FBS for 20 h, followed by executioner caspase activity measurement using the Quantipak kit (Biomol). This was performed on 50 μ g cytosolic cell extracts using 200 μ M of the caspase-3 chromogenic substrate Z-Asp-Glu-Val-Asp-*para*-nitroaniline (DEVD-pNA), according to the manufacturer's instructions. Samples containing no cell extract were used as negative controls.

Time-lapse microscopy. To study cell migration, CECs were seeded on 35-mm dishes in DMEM growth medium, followed 1 h later by filming of the cells for 12 h at 37 °C. Images were collected with a KPM1E/K-S10 CCD camera (Hitachi Denshi) every 10 min using Tempus software (Kinetic Imaging Ltd, now Andor Technology). Time-lapse images from the whole duration of the experiment were then displayed as a movie, and cells from each frame tracked using Motion Analysis software (Andor/Kinetic Imaging). The mean migration speed of the population was determined using Mathematica 6.0 workbooks³⁰.

Scratch wound assay. iCECs or HUVECs 24 h after transfection were seeded at high density in DMEM growth medium and EBM-2 medium, respectively, followed one day later by wounding of the confluent monolayer by scraping a pipette tip across.

siRNA knockdown experiments. HUVECs (passage 2; Lonza) were plated in six-well dishes at 1×10^5 cells per well, 24 h before transfection. Four siRNA oligonucleotides (Dharmacon) targeting the human p110 α catalytic subunit, were used individually or in combination: D-003018-5 (sense sequence (GCUA-UCAUUGAACAAUUAUU), antisense sequence (5'P UAAUUGUUCAGAUAGUAGCUU)), D-003018-6 (sense sequence (GGAUAGAGGCCAAUUAUUAUU), antisense sequence (5'P UAUUAUUUGGCCUCUAUCCUU)), D-003018-7 (sense sequence (GGACAACUGUUUUAUUAUUAUU), antisense sequence (5'P CUUAUUGAAACAGUUGUCCUU)) and D-003018-8 (sense sequence (GCCAGUACCUCAUGGAUUAUU), antisense sequence (5'P UAAUCCAUGAGGUACUGGCUU)).

The siRNA pool gave the strongest knockdown (data not shown) and was used in all subsequent experiments. Cells were transfected with Oligofectamine (Invitrogen), and after 24 h cells were trypsinized and plated at sub-confluence and allowed to adhere (8 h) on fibronectin-coated coverslips for immunofluorescence analysis (5×10^4 cells per well) or at confluence in 24-well plates for time-lapse microscopy or lysis and western blot analysis (10^5 cells per well). Cell proliferation was analysed by Casy Counter analysis 24, 48 and 72 h after transfection. For cell viability analysis, cells were trypsinized 48 h after transfection, followed by staining with propidium iodide. Samples were analysed by flow cytometry.

Retinal angiogenesis. Postnatal P5 eyes ($n \geq 4$ for each genotype) were fixed in 4% PFA in PBS at 4 °C overnight and washed in PBS. Retinas were dissected, stained with biotinylated isolectin-B4 and flat-mounted as previously described or processed for multiple labelling²¹.

Endothelial cell proliferation in vivo. BrdU was injected at 50 mg kg⁻¹ body weight intraperitoneally into P5 pups 2 h before harvesting of the eyes. P5 eyes were fixed overnight in 4% PFA in PBS at 4 °C and washed in PBS. Retinas were dissected, incubated with proteinase K (10 μ g ml⁻¹ in PBT) for 30 min, followed by 10 min in glycine (2 mg ml⁻¹ in PBT), two washes in PBT and post-fixation for 20 min in 4% PFA/0.1% glutaraldehyde in PBT. After washing in PBT, samples were equilibrated in DNase I buffer (50 mM Tris.HCl, 10 mM MgCl₂

at pH 7.5) and incubated for 2 h in DNase I ($0.1 \text{ U } \mu\text{l}^{-1}$) at 37°C . DNase was heat-inactivated for 10 min at 70°C in 50 mM Tris-HCl (pH 7.5), followed by washes in ice-cold Tris-HCl. After two washes in PBT, unspecific binding of antibodies was blocked by 1% BSA/0.5% Tween 20 in PBS for 30 min. Samples were labelled using isolectin-B4 and mouse anti-BrdU as described above. Samples were flat-mounted and cover-slipped using Moviol.

Hindbrain angiogenesis. E11.5 hindbrains ($n \geq 5$ for each genotype), dissected clean from surrounding tissue, were fixed in 4% formaldehyde in PBS for 2 h on ice, permeabilized by two washes in PBS containing 0.1% Triton X-100 (PBX), incubated for 5 min in 50% methanol/PBX then incubated in methanol. Samples were rehydrated for 30 min at room temperature, stained with CD31 and flat-mounted as previously described. Littermate samples were photographed at the same magnification. For each genotype, the average number of branch points was determined in five randomly chosen $500 \mu\text{m}^2$ regions²².

BrdU incorporation. E11.5 timed pregnant females were injected intraperitoneally with 50 mg kg^{-1} body weight BrdU in PBS and killed 2 h later. Dissected embryonic hindbrains were processed for immunostaining as described above, except they were incubated for 30 min in 4 N HCl, followed by neutralization in 100 mM sodium tetraborate for 5 min before the staining procedure. Samples were labelled using isolectin-B4 and mouse anti-BrdU as described above. Samples were flat-mounted and cover-slipped using Moviol. The analysis was performed in four or five hindbrains per genotype.

Measurement of Rho activity. The capacity of Rac1-GTP to bind GST-PBD (p21-activated kinase-binding domain), or RhoB-GTP and RhoA-GTP to bind to GST-rhotekin-RBD, was used to determine the cellular levels of active GTPases. Rho/Rac-GTP pull-down assays were performed by incubation of cleared lysates with glutathione-Sepharose 4B beads coupled to GST-PBD or GST-rhotekin-RBD (Upstate Biotechnology) for 1 h at 4°C , followed by four washes of the beads in extraction buffer (50 mM Tris-HCl pH 7.5, 1 mM EDTA, 500 mM NaCl, 10 mM MgCl_2 , 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 10% glycerol, 0.5% 2-mercaptoethanol). Bound proteins were solubilized by the addition of 25 μl of SDS-polyacrylamide gel electrophoresis (SDS-PAGE) Laemmli loading buffer, followed by separation on 12.5% SDS-PAGE gels and western blotting for Rac, RhoB or RhoA³¹.

Immunofluorescence. Cells grown on coverslips were fixed with 4% paraformaldehyde for 20 min, blocked with TBS containing 1% BSA (25 mM Tris-HCl pH 7.4, 150 mM NaCl) for 10 min, permeabilized for 5 min with PBS containing 0.2% Triton X-100 at 4°C , and incubated at 37°C for 30 min with antibodies or $1 \mu\text{g ml}^{-1}$ TRITC-labelled phalloidin. Specimens were mounted in DAKO fluorescent mounting medium (Ely). Confocal laser scanning microscopy was performed with an LSM 510 (Zeiss) mounted over an Axioplan microscope (Zeiss) using a $\times 40$ 1.3 NA oil immersion objective.

Transient transfection. p110 $\alpha^{\text{D933A/WT}}$ iCECs were transfected using Lipofectamine 2000 (GIBCO) according to the manufacturer's instructions. Plasmids used were pCMV5-myc-Rac1-V12, pCMV5-Flag-RhoA-V14 and pCMV5-HA-RhoB-V14 or empty vector (pCMV-Myc; Clontech, catalogue number K6003-1). Twenty-four hours after transfection, the cells were trypsinized and seeded on Matrigel for analysis of tube formation as described above.

Statistical analysis. Means were compared between two groups using the non-parametric Mann-Whitney *U*-test. For a three-group comparison, the analysis of variance (ANOVA) test was used.

26. Bilancio, A. *et al.* Key role of the p110delta isoform of PI3K in B-cell antigen and IL-4 receptor signaling: comparative analysis of genetic and pharmacologic interference with p110 δ function in B cells. *Blood* **107**, 642–650 (2006).
27. Lidington, E. A. *et al.* Conditional immortalization of growth factor-responsive cardiac endothelial cells from H-2K(b)-tsA58 mice. *Am. J. Physiol. Cell Physiol.* **282**, C67–C74 (2002).
28. Weis, S. *et al.* Src blockade stabilizes a Flk/cadherin complex, reducing edema and tissue injury following myocardial infarction. *J. Clin. Invest.* **113**, 885–894 (2004).
29. Cutillas, P. R. *et al.* Ultrasensitive and absolute quantification of the phosphoinositide 3-kinase/Akt signal transduction pathway by mass spectrometry. *Proc. Natl Acad. Sci. USA* **103**, 8959–8964 (2006).
30. Wells, C. M. & Ridley, A. J. Analysis of cell migration using the Dunn chemotaxis chamber and time-lapse microscopy. *Methods Mol. Biol.* **294**, 31–41 (2005).
31. Ren, X. D. & Schwartz, M. A. Determination of GTP loading on Rho. *Methods Enzymol.* **325**, 264–272 (2000).

Rapid cloning of high-affinity human monoclonal antibodies against influenza virus

Jens Wrammert^{1*}, Kenneth Smith^{3*}, Joe Miller¹, William A. Langley¹, Kenneth Kokko², Christian Larsen², Nai-Ying Zheng³, Israel Mays³, Lori Garman³, Christina Helms³, Judith James^{4,6,7}, Gillian M. Air⁸, J. Donald Capra^{5,9}, Rafi Ahmed¹ & Patrick C. Wilson^{3,6,9}

Pre-existing neutralizing antibody provides the first line of defence against pathogens in general. For influenza virus, annual vaccinations are given to maintain protective levels of antibody against the currently circulating strains. Here we report that after booster vaccination there was a rapid and robust influenza-specific IgG⁺ antibody-secreting plasma cell (ASC) response that peaked at approximately day 7 and accounted for up to 6% of peripheral blood B cells. These ASCs could be distinguished from influenza-specific IgG⁺ memory B cells that peaked 14–21 days after vaccination and averaged 1% of all B cells. Importantly, as much as 80% of ASCs purified at the peak of the response were influenza specific. This ASC response was characterized by a highly restricted B-cell receptor (BCR) repertoire that in some donors was dominated by only a few B-cell clones. This pauciclonal response, however, showed extensive intraclonal diversification from accumulated somatic mutations. We used the immunoglobulin variable regions isolated from sorted single ASCs to produce over 50 human monoclonal antibodies (mAbs) that bound to the three influenza vaccine strains with high affinity. This strategy demonstrates that we can generate multiple high-affinity mAbs from humans within a month after vaccination. The panel of influenza-virus-specific human mAbs allowed us to address the issue of original antigenic sin (OAS): the phenomenon where the induced antibody shows higher affinity to a previously encountered influenza virus strain compared with the virus strain present in the vaccine¹. However, we found that most of the influenza-virus-specific mAbs showed the highest affinity for the current vaccine strain. Thus, OAS does not seem to be a common occurrence in normal, healthy adults receiving influenza vaccination.

Influenza causes 36,000 deaths annually in the USA alone, and the influenza pandemic of 1918 caused an estimated 50 million deaths worldwide². Outbreaks of avian influenza infections in human populations that caused substantially higher mortality rates foresee the possibility of another deadly pandemic³. The challenge of influenza has long been to design vaccines that induce long-lasting immunity against a pathogen that rapidly alters its appearance to the immune system by mutating (antigenic drift) and exchanging (antigenic shift) its components. Antibodies play a key role in protection against influenza infection^{4–7}. However, the underlying B-cell response leading to the rapid production of ASCs that secrete antibodies is only

beginning to be understood^{8–12}. Critically, we do not yet know if B-cell memory can provide sufficient protection early in the response to counteract variant strains of influenza or if the response is dominated by antibodies previously generated against divergent viruses in an OAS fashion. Finally, of profound clinical significance is the possibility that the early ASC response observed after immunization can be exploited to generate therapeutic or diagnostic mAbs rapidly to emerging influenza virus strains, or to any immunizing antigen.

To determine the dynamics and magnitude of the human anti-influenza response, we analysed the frequency of ASCs and memory B cells in a time course after vaccination. The ASC response was quite transient, peaking at approximately day 7 and returning to barely detectable levels by day 14 after vaccination (Fig. 1a, b). The frequency of influenza-specific ASCs averaged 6.4% (about 2,500 ASCs per millilitre of blood) at day 7, and accounted for up to 16% of all B cells (range for ten donors: 1.1–16%, Fig. 1b). Also, most of these ASCs were generated during the vaccination response as almost all ASCs expressed the protein bound by the Ki-67 antibody, indicating recent proliferation, and most expressed homogeneously high levels of human leukocyte antigen DR (HLA-DR)¹³ (Fig. 1c). Importantly, analysis of IgG-secreting ASCs isolated by cell sorting at day 7 post-immunization demonstrated that most were influenza-vaccine-specific (ranging from 20% to 85%, average 70%; Fig. 1d). The ASCs were mainly IgG positive, with minor components of IgA- and IgM-positive cells (data not shown), suggesting an origin from the memory B-cell compartment. The memory B-cell response was also quantified¹⁴. Increasing from low levels before vaccination, influenza-specific memory B cells peaked a week after the ASC response at 14–28 days after vaccination and averaged 8.2% of the IgG⁺ memory B cells, or about 1% of all B cells (Fig. 1e). We conclude that influenza vaccination results in a massive burst of IgG⁺ ASCs that are predominantly influenza reactive and peak at approximately day 7 post-immunization.

The rapid accumulation of ASCs suggests that the response could be highly clonal in nature, limiting the early influenza response. Some clonal activation of ASCs occurs after tetanus vaccination¹². We therefore analysed the immunoglobulin repertoire breadth (that is, the variable genes and junctional diversity) of the influenza-specific ASCs. Influenza vaccination caused a surprisingly pauciclonal response, with some donors being dominated by the progeny of only a few expanded B-cell clones (Fig. 2a and Supplementary Fig. 1a).

¹Emory Vaccine Center and Department of Microbiology and Immunology, Emory University School of Medicine, Atlanta, Georgia 30322, USA. ²Emory Transplant Center and Department of Surgery, Emory University School of Medicine, Atlanta, Georgia 30322, USA. ³Immunobiology and Cancer Research Program, The Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma 73104, USA. ⁴Arthritis and Immunology Research Program, The Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma 73104, USA. ⁵Molecular Immunogenetics Research Program, The Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma 73104, USA. ⁶Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73104, USA. ⁷Department of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73104, USA. ⁸Department of Biochemistry and Molecular Biology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73104, USA. ⁹Department of Microbiology and Immunology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73104, USA.

*These authors contributed equally to this work.

Clonal expansions accounted for 43% of the ASC variable regions from the 14 immunized donors, including three with over 70% clonality (Fig. 2a). In stark contrast, based on VH regions sequenced from our laboratory in a comparable fashion^{15–17}, naive and memory B cells (IgM or IgG) isolated from blood were rarely or never clonal, whereas for tonsillar B cells only 10% of IgM and 12% of IgG germinal centre and memory cells were clonally related.

Immunoglobulin variable region somatic hypermutation allows for the generation of high-affinity antibodies^{18,19}. Surprisingly, the influenza-specific ASCs had accumulated more somatic mutations than any normal population of B cells. Considering the various donors (Fig. 2b), the ASCs averaged 19.4 ± 3.5 VH gene mutations, which is greater than that of germinal centre or memory B cells which average 13.6 ± 4.8 mutations for IgG or 8.4 ± 3.8 mutations for IgM. A surprising 11% (41/405) of the ASC VH gene segments have more

than 30 of 300 (or about 10%) of the total nucleotides altered (Fig. 2c). A preference for complementarity-determining region replacement mutations suggests that the ASCs were functionally selected (Supplementary Table 1). These observations suggest the origin of the anti-influenza ASCs is predominantly memory B cells that probably accumulated new mutations on this and on previous rounds of activation.

It is not known how often the ASCs that are induced by vaccination produce high-affinity antibodies against influenza. Immunoglobulin variable region genes from ASCs can be used to express specific antibodies²⁰. We therefore used the variable gene transcripts of isolated single ASCs to express recombinant mAbs in the human 293 cell line (Supplementary Fig. 1b and Methods). The ASCs of five donors were isolated at day 7 post-vaccination. From these, 71% (61/86) of the antibodies bound with high affinity to either native antigens of

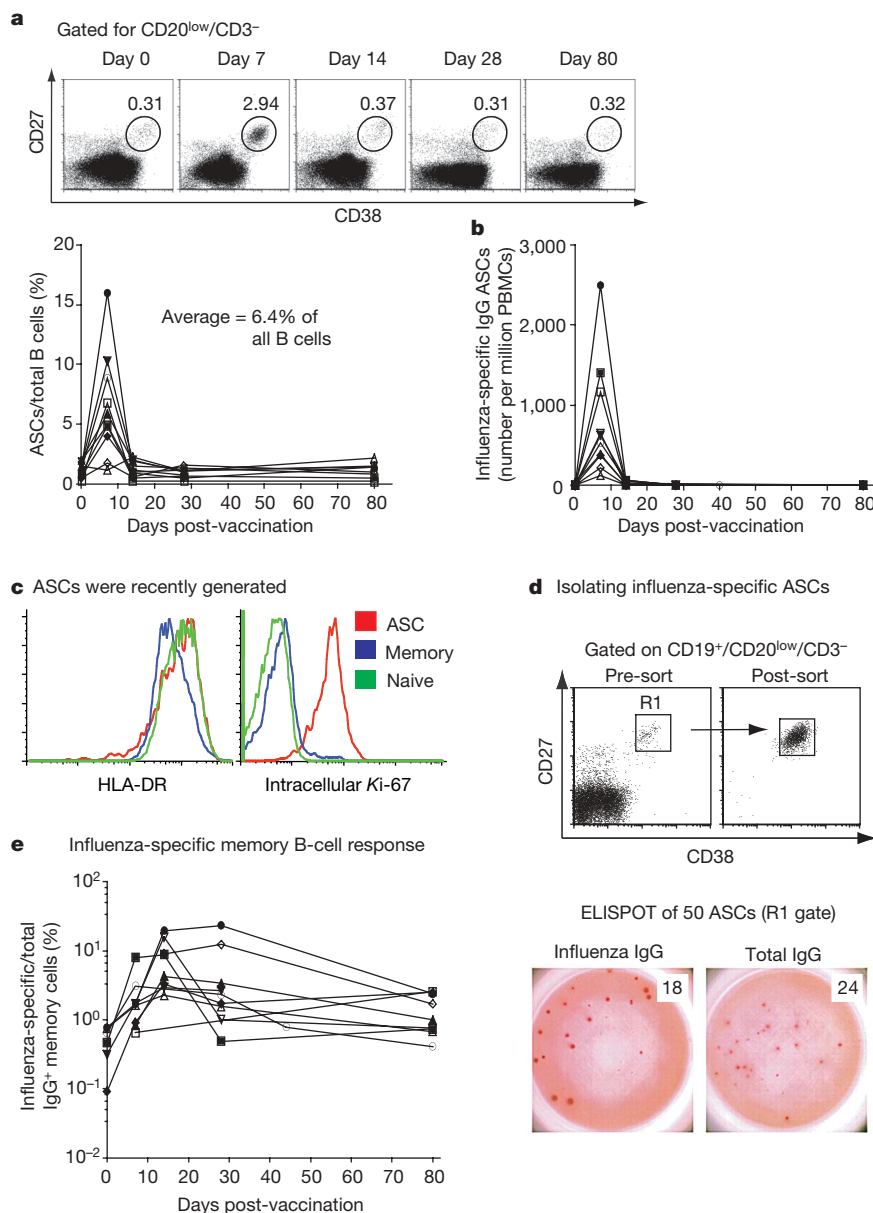


Figure 1 | Analysis of the B-cell response induced by influenza vaccination. **a**, ASCs were measured in blood by flow cytometry. Shown is the frequency of the ASC gate (CD3⁻/CD20^{low}/CD19⁺/CD27^{hi}/CD38^{hi}) for a representative donor and a summary for all ten donors normalized to total CD20⁺/CD19⁺ B-cell numbers. **b**, Peripheral blood mononuclear cells (PBMCs) collected from ten donors were assayed for influenza-specific IgG secreting ASCs by ELISPOT assay at 0, 7, 14, 28 and 80 days after vaccination. Each sample was measured in duplicate, averaged and plotted as

ASCs per million PBMCs over time post-vaccination. **c**, HLA-DR and intracellular expression of Ki-67 by ASCs compared with naive or memory B cells. **d**, Most ASCs at day 7 after influenza vaccination are influenza specific. Influenza- and total IgG-specific ELISPOT assays from several donors gave similar results. **e**, Percentage of influenza-specific memory cells per total IgG-positive memory cells after mitogen stimulation as measured by ELISPOT at 0, 7, 14, 28 and 80 days post-vaccination as previously described¹⁴.

the influenza vaccine strains (53/86, 61%) or to components of the vaccine only (8/86, 9%) (Fig. 3, Table 1 and Supplementary Fig. 2). We suspect that the epitopes found only in the vaccine are exposed on the fixed virions or are from added preservatives. In comparison, none of the 86 mAbs generated from naive B cells¹⁵ (Fig. 3b) and only one of 54 antibodies from random IgG memory B cells bound to the influenza vaccine strains with appreciable affinity (data not shown). The antibodies produced from the influenza-specific ASCs bound to any of the three vaccine components with similar frequency (Table 1 and Supplementary Fig. 2). Analysis of viral antigen specificity by immunoprecipitation and western blot (Supplementary Fig. 3) found that 60% of the influenza-reactive antibodies bound to haemagglutinin, of which half were haemagglutination inhibiting (HAI) (Fig. 3c and Table 1). Twelve per cent of the antibodies bound to neuraminidase or to other minor components of the vaccine likely residual to the purification of haemagglutinin and neuraminidase during vaccine production. Ten per cent of the antibodies did not precipitate native antigens and bound only to epitopes on denatured viral proteins detectable by western blot. Importantly, each of three representative HAI⁺ antibodies against influenza-A (anti-H3N2) and one against influenza-B from the day-7 ASCs (Table 1, bold) were found to neutralize viral infection of Madin-Darby canine kidney (MDCK) cells *in vitro* (each neutralized virus at less than $1 \mu\text{g ml}^{-1}$ antibody, Supplementary Fig. 4). In conclusion, after

influenza vaccination, early ASCs produce functional antibodies that bind with high affinity and likely provide early protection.

Although most of the ASCs arise only after vaccination (Fig. 1a), 29% of the antibodies generated did not detectably bind to the

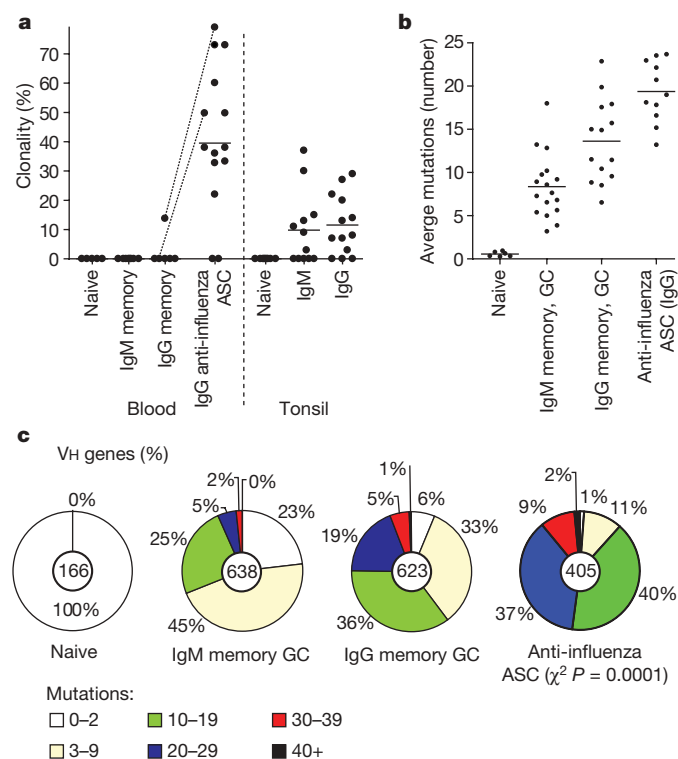


Figure 2 | The ASC response after influenza vaccination is pauci-clonal and highly diversified by somatic hypermutation. **a**, Comparison of the mean proportion (line) of all clonal variable region sequences from day-7 ASCs of 14 donors (points), including the bulk RNA of 10^4 – 10^5 ASCs from ten donors and verification by single-cell RT-PCR for four donors (average 37 sequences per donor). The ASCs were the most clonally related population (*t*-test, $P \leq 0.0003$). Dotted lines indicate donors from which memory and ASCs were analysed simultaneously. Other B-cell populations were from historical data analysed in a similar fashion from our laboratory^{15–17} (see Methods and Supplementary Data) **b**, Each point is the average frequency of somatic mutations per sequence from each donor (*n* values within Methods). On average, the anti-influenza ASCs had accumulated more mutations than either the IgG (*t*-test $P = 0.003$) or IgM ($P = < 0.0001$) memory and germinal-centre populations. GC, germinal centre. **c**, The proportion of all variable genes from each B-cell population with the number of somatic mutations denoted in the legend (*n* values are at the centre of each pie chart).

a 71% (61 of 86) of IgG ASCs bind influenza virions (62%, 53/86) or the vaccine only (9%, 8/86, not shown) ($K_d = 10^{-8}$ – 10^{-11})

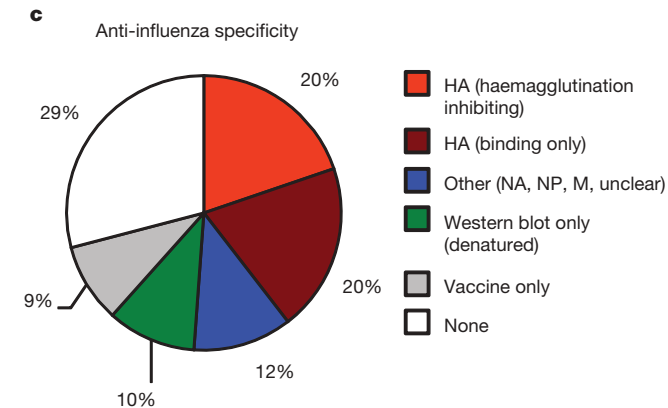
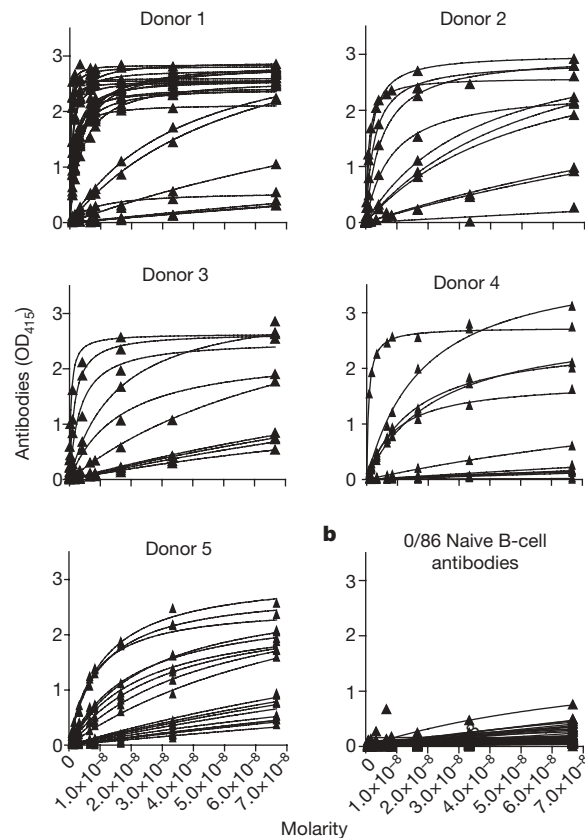


Figure 3 | High-affinity mAbs generated from single influenza-specific ASCs. **a**, Recombinant mAbs from day-7 IgG anti-influenza ASCs (Supplementary Fig. 1b) bind to a mixture of the three influenza vaccine strain virions with high affinity. In total, 71% of the ASC antibodies bound either native antigens of influenza viruses freshly grown in eggs (53/86, 62%), or to antigens within the vaccine only (8/86, 9%, not shown). Antibodies from each of the five donors were influenza specific (by donor, 34, 13, 11, 15 and 21 antibodies were generated, of which 45–85% were influenza specific). Individual antibody strain specificities are shown in Table 1 and Supplementary Fig. 2. **b**, None of 86 naive B-cell antibodies bound influenza. **c**, Analysis by immunoprecipitation and western blot (Supplementary Fig. 3) identified the specific viral antigens bound. Haemagglutination assays identified those antibodies that were inhibiting (Table 1 and Methods). HA, haemagglutinin; NA, neuraminidase; NP, nuclear protein; M, matrix protein.

influenza strains or whole vaccine (Fig. 3c). Possible causes include errors introduced by the reverse-transcription polymerase chain reaction (RT-PCR) steps (though PCR errors were rare, Supplementary Table 2), targeting of non-viral or denatured components of the vaccine or antigens only evident physiologically, bystander activation of non-specific memory cells⁸, or displacement of non-specific plasma cells from the bone marrow¹³. The last possibility is unlikely as expression of HLA-DR¹³ and Ki-67 (Fig. 1c) by the ASCs suggests they were newly generated.

The long-held theory of OAS suggests that new influenza variants will evade surveillance when memory B cells reactive to previous viral strains dominate the response¹. To consider the impact of OAS directly, we compared the relative affinity with either the current B strain virus (B/Malaysia/2506/2004) or with the two previous ones (B/Shanghai/361/2002 or B/Hong Kong/33/2001) (Fig. 4a, b). In the 2006/7 season, antibodies were analysed from five donors who had also been vaccinated in the 2005/6 season and one in 1991, so that

Table 1 | Characteristics of anti-influenza antibodies

H1N1: A/New Caledonia/20/99

| | Antibody | K_d (M) | HAI | Antigen |
|---------|----------------------|------------------------|------|--|
| Donor 1 | D1-1 | 2.70×10^{-10} | None | Haemagglutinin |
| | D1-2 ($\times 2$)* | 7.33×10^{-10} | None | Haemagglutinin |
| | D1-3 ($\times 2$)* | 1.55×10^{-9} | None | Haemagglutinin |
| | D1-4 | 2.14×10^{-9} | None | Haemagglutinin |
| | D1-5 | 2.66×10^{-9} | 1 | Haemagglutinin |
| | D1-6 | 3.62×10^{-9} | 1 | Haemagglutinin |
| Donor 2 | D2-1 | 2.73×10^{-9} | None | 85 kilodalton (kDa) band on western blot |
| | D2-2 | 5.42×10^{-9} | None | 85 kDa band on western blot |
| | D2-3 | 5.20×10^{-9} | None | Haemagglutinin |
| Donor 3 | D3-1 | 1.88×10^{-9} | None | 85 kDa band on western blot |
| Donor 4 | D4-1 | 4.00×10^{-8} | None | 85 kDa band on western blot |
| | D4-2 | 1.65×10^{-8} | None | Haemagglutinin |
| Donor 5 | D5-1 | 5.01×10^{-11} | None | Multiple bands on western blot |
| | D5-2 | 1.01×10^{-9} | None | Denatured haemagglutinin-1 on western blot |
| | D5-3 | 1.78×10^{-8} | None | Haemagglutinin |
| | D5-10 | 1.78×10^{-8} | None | Multiple bands on western blot |

H3N2: A/Wisconsin/67/2005(2006/7) or A/California/7/2004 (for donor 3)

| | | | | |
|---------|--------------------------------------|------------------------|------|-----------------|
| Donor 1 | D1-7 | 7.72×10^{-11} | 128 | Haemagglutinin |
| | D1-8 | 2.86×10^{-10} | 4 | Haemagglutinin |
| | D1-9 ($\times 4$)* | 3.77×10^{-10} | 8 | Haemagglutinin |
| | D1-10 | 4.18×10^{-10} | 4 | Haemagglutinin |
| | D1-11 | 1.57×10^{-9} | None | Nuclear protein |
| Donor 2 | D2-4 | 3.62×10^{-10} | 2 | Haemagglutinin |
| | D2-5 | 8.29×10^{-9} | None | Unclear |
| | D3-2 | 3.50×10^{-9} | None | Haemagglutinin |
| Donor 3 | D3-3 | 1.56×10^{-8} | None | Haemagglutinin |
| | D3-4 | 4.86×10^{-10} | 32 | Haemagglutinin |
| Donor 4 | D4-3 | 4.56×10^{-9} | 1 | Haemagglutinin |

B strains: B/Malaysia/2506/2004 or B/Shanghai/361/2004 (donors 3 and 6)

| | | | | |
|---------|-----------------------|------------------------|------|---|
| Donor 1 | D1-12 | 1.93×10^{-10} | None | Haemagglutinin |
| | D1-13 | 2.04×10^{-10} | 16 | Haemagglutinin |
| | D1-14 ($\times 2$)* | 2.43×10^{-10} | None | Haemagglutinin |
| | D1-15 | 2.47×10^{-10} | 8 | Haemagglutinin |
| | D1-16 | 6.20×10^{-10} | None | Haemagglutinin |
| | D1-17 | 6.33×10^{-10} | None | Haemagglutinin |
| | D1-18 | 4.74×10^{-10} | None | Haemagglutinin |
| | D1-19 | 3.70×10^{-8} | None | Matrix protein on western blot |
| | D2-6 | 1.71×10^{-9} | None | Unclear |
| Donor 2 | D2-7 | 5.77×10^{-8} | None | Matrix protein on western blot |
| | D2-8 | 3.66×10^{-8} | None | Haemagglutinin |
| | D3-5 | 1.66×10^{-8} | None | Unclear (low affinity) |
| Donor 3 | D3-6 | 1.62×10^{-8} | None | Denatured nuclear protein on western blot |
| | D4-4 | 1.71×10^{-8} | 1 | Haemagglutinin |
| Donor 4 | D4-5 | 3.28×10^{-8} | 4 | Haemagglutinin |
| | D5-5 | 3.78×10^{-8} | None | Unclear (low affinity) |
| Donor 5 | D5-6 | 1.11×10^{-8} | None | Matrix protein on western blot |
| | D5-7 | 1.26×10^{-8} | None | Unclear (low affinity) |
| | D5-8 | 4.8×10^{-8} | None | 85 kDa band on western blot |
| | D6-1 | 5.04×10^{-10} | 4 | Haemagglutinin |

* Clonal expansions with number of clones indicated; bold, mAbs tested for viral neutralization.

reactive memory cells should have been readily available for an OAS response. Importantly, each of the 19 anti-B strain antibodies bound to the new B strain with equal, and in most cases with greater, affinity than the previous vaccine strains (Fig. 4c and Supplementary Fig. 2). This adaptation occurred despite the 10% or less difference of the

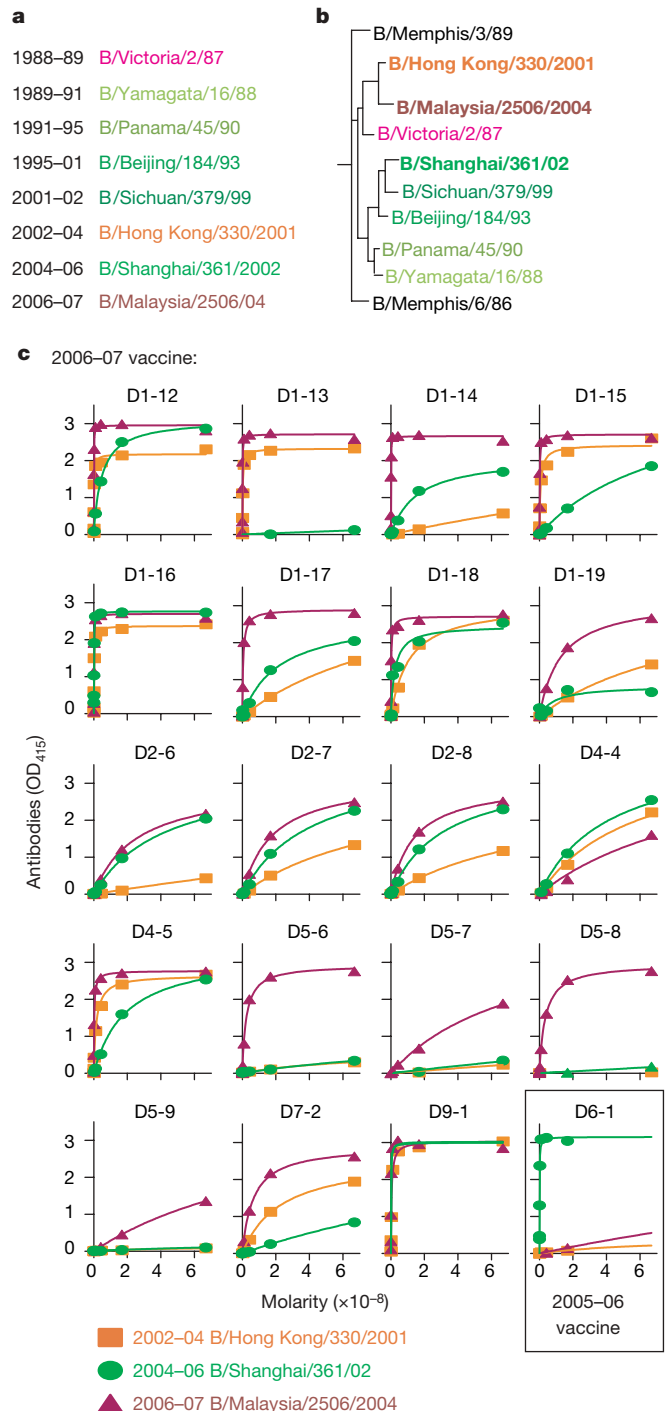


Figure 4 | Specificity for the newly introduced influenza-B strain in the vaccine suggests a minimal impact of OAS. **a**, Influenza-B strains used for the vaccine since 1989. Throughout the figure, strain names are colour coded for the Yamagata lineage (green) and the Victoria lineage (orange/red). **b**, Phylogenetic tree illustrating the similarity of recent influenza-B strains and the years in which each strain was included in the vaccine. The three vaccine strains tested (bold font) included B/Malaysia/2506/2004 (2006/7 season), which is most similar to the 2002–2004 strain (B/Hong Kong/33/2001). Conversely, the 2005/6 vaccine strain, B/Shanghai/361/2002, is more divergent. **c**, All anti-B-strain antibodies reacted with equal or greater affinity to the current year's vaccines when tested by ELISA.

haemagglutinin sequence of the 2006/7 B strain from those used in previous vaccines. Although previous exposure to B/Malaysia/2506/2004 cannot be entirely excluded, there was no history of exposure, and pre-vaccination serum titres of antibody against B/Malaysia/2506/2004 were not above background levels (data not shown). Thus we conclude that even for the earliest detectable influenza-specific B cells after vaccination, the ASCs, OAS does not limit reactivity to newly introduced influenza strains.

In conclusion, we show that after influenza vaccination we can isolate an almost entirely antigen-specific population of ASCs that comprise about 5% of all blood-borne B cells. Further, our findings help to resolve a major, long-standing obstacle in the field of medicine²¹: the rapid production of fully human mAbs. Antibody or serum therapy has been demonstrated to treat a plethora of diseases effectively, but it is not widely used because sometimes fatal anaphylactic responses and serum sickness are common. These obstacles can only be overcome by using fully human mAbs. Our findings demonstrate that we can now generate human mAbs from the antigen-specific ASCs directly, and within only weeks of vaccination (Supplementary Fig. 1c). With a modern resurgence of interest in monoclonal antibody therapy, we anticipate that antibodies produced from post-vaccination ASCs will generate substantial advances for the treatment of infectious diseases.

Conventional wisdom holds that the level of pre-formed antibody is the main correlate of protection against influenza virus. However, our results, showing the rapidity of the antibody response after vaccination and the high affinity of the antibodies produced, strongly suggest that the recall response could also play a role in protective immunity. This antibody would not, of course, prevent initial infection but could play a crucial role in preventing the spread of virus and bringing about faster resolution of the infection. This notion is supported by our finding that OAS was not a significant aspect of the memory response, as the antibodies produced were highly specific to the immunizing antigen.

METHODS SUMMARY

Detailed methods are found in the Methods and Supplementary Information. Healthy volunteers received either the Fluzone 2005-06 (Aventis Pasteur Inc) or Fluvirin 2006-07 (Chiron) influenza vaccine formulations. This project had institutional review board approvals from both Emory University School of Medicine and the Oklahoma Medical Research Foundation. Enumeration of IgG or influenza-specific B cells has been previously described, including: ELISPOT²² and polyclonal activation of memory B cells^{14,22}. HAI titres and viral neutralization were determined by standard procedures²³. Monoclonal antibodies were tested for HAI at 30 µg ml⁻¹ and five twofold serial dilutions. The ASCs were identified as CD3⁺/CD20^{-low}/CD19⁺/CD27^{hi}/CD38^{hi} cells. Isolation of other B-cell types and methods for RT-PCR of the variable genes were as previously described^{15,17}. An average of 35 variable gene sequences was analysed from each donor (Methods and Supplementary Information). The leader and constant regions that are rarely targeted by somatic mutation were primed for RT and PCR to avoid biases between populations for the variable gene repertoire and clonality determinations. The single-cell RT-PCR methods and the procedures for production of recombinant mAbs were as previously described^{15,24}. Monoclonal antibodies were screened against fresh influenza virions grown in chicken eggs. Antibody affinities (K_d) were calculated by nonlinear regression (GraphPad Prism software) of influenza enzyme-linked immunosorbent assay (ELISA) curves plotted from eight dilutions of antibody ranging from 10 µg ml⁻¹ to 0.125 µg ml⁻¹.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 16 October 2007; accepted 4 March 2008.

Published online 30 April 2008.

- Francis, T. Jr. On the doctrine of original antigenic sin. *Proc. Am. Phil. Soc.* **104**, 572–578 (1960).

- Ahmed, R., Oldstone, M. B. & Palese, P. Protective immunity and susceptibility to infectious diseases: lessons from the 1918 influenza pandemic. *Nature Immunol.* **8**, 1188–1193 (2007).
- Subbarao, K. & Joseph, T. Scientific barriers to developing vaccines against avian influenza viruses. *Nature Rev. Immunol.* **7**, 267–278 (2007).
- Gerhard, W. *et al.* Role of the B-cell response in recovery of mice from primary influenza virus infection. *Immunol. Rev.* **159**, 95–103 (1997).
- Luke, T. C., Kilbane, E. M., Jackson, J. L. & Hoffman, S. L. Meta-analysis: convalescent blood products for Spanish influenza pneumonia: a future H5N1 treatment? *Ann. Intern. Med.* **145**, 599–609 (2006).
- Puck, J. M., Glezen, W. P., Frank, A. L. & Six, H. R. Protection of infants from infection with influenza A virus by transplacentally acquired antibody. *J. Infect. Dis.* **142**, 844–849 (1980).
- Simmons, C. P. *et al.* Prophylactic and therapeutic efficacy of human monoclonal antibodies against H5N1 influenza. *PLoS Med.* **4**, e178 (2007).
- Bernasconi, N. L., Traggiai, E. & Lanzavecchia, A. Maintenance of serological memory by polyclonal activation of human memory B cells. *Science* **298**, 2199–2202 (2002).
- Brokstad, K. A. *et al.* Parenteral influenza vaccination induces a rapid systemic and local immune response. *J. Infect. Dis.* **171**, 198–203 (1995).
- Brokstad, K. A. *et al.* Parenteral vaccination against influenza does not induce a local antigen-specific immune response in the nasal mucosa. *J. Infect. Dis.* **185**, 878–884 (2002).
- Sasaki, S. *et al.* Comparison of the influenza virus-specific effector and memory B-cell responses to immunization of children and adults with live attenuated or inactivated influenza virus vaccines. *J. Virol.* **81**, 215–228 (2007).
- Poulsen, T. R. *et al.* Kinetic, affinity, and diversity limits of human polyclonal antibody responses against tetanus toxoid. *J. Immunol.* **179**, 3841–3850 (2007).
- Odendahl, M. *et al.* Generation of migratory antigen-specific plasma blasts and mobilization of resident plasma cells in a secondary immune response. *Blood* **105**, 1614–1621 (2005).
- Crotty, S., Aubert, R. D., Glidewell, J. & Ahmed, R. Tracking human antigen-specific memory B cells: a sensitive and generalized ELISPOT system. *J. Immunol. Methods* **286**, 111–122 (2004).
- Koelsch, K. *et al.* Mature B cells class switched to IgD are autoreactive in healthy individuals. *J. Clin. Invest.* **117**, 1558–1565 (2007).
- Zheng, N. Y., Wilson, K., Jared, M. & Wilson, P. C. Intricate targeting of immunoglobulin somatic hypermutation maximizes the efficiency of affinity maturation. *J. Exp. Med.* **201**, 1467–1478 (2005).
- Zheng, N. Y. *et al.* Human immunoglobulin selection associated with class switch and possible tolerogenic origins for C delta class-switched B cells. *J. Clin. Invest.* **113**, 1188–1201 (2004).
- Clarke, S. H. *et al.* Inter- and intraclonal diversity in the antibody response to influenza haemagglutinin. *J. Exp. Med.* **161**, 687–704 (1985).
- Cook, W. D. & Scharff, M. D. Antigen-binding mutants of mouse myeloma cells. *Proc. Natl Acad. Sci. USA* **74**, 5687–5691 (1977).
- Meijer, P. J. *et al.* Isolation of human antibody repertoires with preservation of the natural heavy and light chain pairing. *J. Mol. Biol.* **358**, 764–772 (2006).
- Llewellyn, M. B., Hawkins, R. E. & Russell, S. J. Discovery of antibodies. *Br. Med. J.* **305**, 1269–1272 (1992).
- Crotty, S. *et al.* SAP is required for generating long-term humoral immunity. *Nature* **421**, 282–287 (2003).
- Webster, R., Cox, N. & Stohr, K. *World Health Organization Manual on Animal Influenza Diagnosis and Surveillance* (WHO, Geneva, 2002).
- Wardemann, H. *et al.* Predominant autoantibody production by early human B cell precursors. *Science* **301**, 1374–1377 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Popkowski, H. Wu, L. Abraham and B. Begley for technical assistance, and R. Casellas and J. Knight for reading the manuscript. This work was funded in parts by National Institutes of Health (NIH) grant numbers HHSN266200500026C (P.C.W.), P20 RR018758 (P.C.W.), NIH/National Institute of Allergy and Infectious Diseases (NIAID) U19-AI057266-04 (R.A.), NIH/NIAID HHSN266200700006C Center of Excellence for Influenza Research and Surveillance (R.A.) and NIH/NIAID N01-AI-50025-02 (R.A. and C.L.). J.W. was supported by a postdoctoral fellowship from The Swedish Research Council.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.C.W. (wilsonp@omrf.org).

METHODS

Cell and serum isolation. All studies were pre-approved by the institutional review boards of Emory University School of Medicine and the Oklahoma Medical Research Foundation. Healthy volunteers received influenza vaccine formulations (Fluzone, Aventis Pasteur, 2005/6, or Fluvirin, Chiron, 2006/7). PBMCs were isolated using Vacutainer tubes (Becton Dickinson) or lymphoprep gradient (CellGro), washed and re-suspended in supplemented culture media or PBS. Plasma was heat inactivated.

ELISPOT and memory B-cell assays. ELISPOT and memory assays were as previously described^{14,25}. Total IgG secreting or influenza-specific ASCs were detected using 1/20 diluted influenza vaccine in PBS (as above) or with goat anti-human Ig (Caltag). Dilutions of washed PBMCs incubated in supplemented RPMI medium for 2 h were incubated in ELISPOT plates for 6 h. After washing the plates, ASC antibody was detected with anti-huIgG-biotin (Caltag) and avidin-D-HRP (Vector Laboratories) and developed with AEC substrate (Sigma) before analysis on an ELISPOT counter (Cellular Technologies Ltd.). Memory cells were detected by incubating PBMCs at 5×10^5 cells per millilitre in R-10 supplemented with pokeweed mitogen extract, phosphothiolated CpG ODN-2006²⁶ and *Staphylococcus Aureus* Cowan (Sigma). After culture for six days, the cells were washed and quantified by ELISPOT.

Flow cytometry and cell sorting. Flow cytometry was performed on whole blood after lysis of erythrocytes. Mostly Pharmingen antibodies were used for quantifying ASC or memory cells (Fig. 1) except anti-CD27-APC (ebiosciences) and goat anti-huIgG-FITC (Southern Biotechnologies). For single-cell analysis and production of mAbs, antibodies used included anti-CD3-FITC, anti-CD20-FITC, anti-CD38-APC-Cy5.5, anti-CD27-PE, anti-IgG-Alexa-647 and anti-CD19-PE-Alexa 610 from Caltag, plus anti-IgD-biotin and streptavidin-PE-Cy7 (Pharmingen). ASCs were gated as $\text{IgG}^+/\text{IgD}^-/\text{CD19}^+/\text{CD3}^-/\text{CD20}^{\text{low}}/\text{CD27}^{\text{high}}/\text{CD38}^{\text{high}}$. All other cell types were isolated as previously described^{15–17}. Cytometry data was analysed using FlowJo software.

Single-cell RT-PCR and PCR of antibody variable region genes. As detailed in Supplementary Information, single B cells were sorted into 96-well PCR plates containing RNase inhibitor (Promega). VH and Vκ genes from each cell were amplified by RT-PCR and nested PCR reactions using cocktails of primers as previously described^{15,24}, then sequenced. To generate recombinant antibodies, restriction sites were incorporated by PCR with primers to the particular variable and junctional genes. RT-PCR of bulk RNA to analyse V genes was as previously described^{15,17,27}.

Analysis of clonality and somatic mutations of variable region genes. To quantify clonality, variable genes were randomly sequenced from the bulk RNA of ASCs from ten donors (by donor, $n = 22, 47, 49, 12, 16, 19, 36, 25, 34$ and 63) and verified by single-cell RT-PCR analysis of ASCs from four donors ($n = 65, 37, 30$ and 50). Naive, memory and germinal-centre cell variable gene libraries included the following VH gene n values: blood naive (by donor, $n = 61, 24, 15, 14$ and 24), blood IgM memory ($n = 28, 17, 27, 11, 23, 12, 29$ and 20), blood IgG memory ($n = 23, 18, 18, 17, 22$ and 21), tonsillar naive B cells ($n = 125, 32, 16, 22, 32, 23, 46$ and 81), tonsillar IgM and germinal centre/memory ($n = 50, 42, 35, 16, 60, 15, 50, 25, 39, 19, 55$ and 58 VH genes) and tonsillar IgG germinal centre/memory ($n = 113, 25, 14, 40, 12, 41, 11, 23, 18, 51, 15, 54$ and 69). The n values for analysis of somatic hypermutation included: anti-influenza ASCs from 11 donors ($n = 63, 18, 33, 46, 49, 11, 36, 11, 30, 35, 25$); IgG germinal centre/

memory cells from 14 donors ($n = 110, 37, 19, 28, 174, 40, 25, 15, 21, 18, 22, 24, 19, 71$); IgM germinal centre/memory from 17 donors ($n = 56, 158, 18, 91, 17, 10, 16, 30, 19, 28, 11, 36, 29, 13, 22, 20, 64$); and naive cells from six donors ($n = 18, 42, 21, 34, 15, 36$). Background mutation rates were insignificant (Supplementary Table 2).

Recombinant monoclonal antibody expression and analysis. All assays are further detailed in the Supplementary Information. VH or Vκ genes amplified from each single cell were cloned into IgG1 or Igκ expression vectors as previously described^{15,24}. Heavy- and light-chain plasmids were co-transfected into the 293A cell line for expression, and antibodies purified with protein A sepharose. The influenza virus strains used for ELISA or HAI were freshly grown in eggs and purified by standard methods²³ and included: A/New Caledonia/20/9 (H1N1), A/California/7/2004 (H3N2) for 2005/6 or A/Wisconsin/67/2005 (H3N2) for 2006/7, and B/Shanghai/361/2002-like for 2006/2007 or B/Malaysia/2506/2004 for 2006/7. Serum or mAb HAI titres were determined as previously described²⁸. After ELISA screening with a cocktail of all influenza strains and 1/20 dilutions of the vaccines, the affinity and specificity of binding-positive mAbs were determined with the individual influenza viruses. ELISA affinities were calculated by nonlinear regression analysis of curves from eight dilutions of antibody ($10\text{--}0.125 \mu\text{g ml}^{-1}$) using GraphPad Prism. Influenza-neutralizing activity was detected as inhibition of MDCK cell death by 50% tissue culture infectious doses of A/Wisconsin/67/2005 or B/Shanghai/361/2002 based on the protocol of the World Health Organization manual²³.

Immunoprecipitation and western blot analyses. All assays are further detailed in the Supplementary Information. For immunoprecipitation, 8 haemagglutinin units (HAU) of virus were lysed and incubated with $10 \mu\text{g ml}^{-1}$ of mAb before purification with Protein A-Sepharose (Pierce). mAb was eluted from the protein-A by boiling in Laemmli buffer (Bio-Rad) and analysed on 12% Tris-glycine polyacrylamide gels. Protein was detected by staining the gels with sypro-orange (1×, Invitrogen). For western blots, 8 HAU of virus was diluted and boiled in denaturing/reducing sample buffer, then run on denaturing polyacrylamide gels (as above) followed by electrophoretic transfer to nitrocellulose membranes. The membranes were incubated with each antibody at $5 \mu\text{g ml}^{-1}$, detected with HRP anti-human IgG (Jackson ImmunoResearch) and developed with ECL plus reagent (GE Healthcare). IP gels and western blot membranes were analysed using a STORM840 system (Molecular Dynamics).

Statistics. Statistical analyses (described in context) were performed using GraphPad Prism: frequencies of clonal relatedness and somatic mutation were compared by non-paired, two-tailed Student's t -tests; χ^2 tests compared summed mutation frequencies.

25. Crotty, S. *et al.* Cutting edge: long-term B cell memory in humans after smallpox vaccination. *J. Immunol.* **171**, 4969–4973 (2003).
26. Hartmann, G. *et al.* Delineation of a CpG phosphorothioate oligodeoxynucleotide for activating primate immune responses *in vitro* and *in vivo*. *J. Immunol.* **164**, 1617–1624 (2000).
27. Wilson, P. C. *et al.* Receptor revision of immunoglobulin heavy chain variable region genes in normal human B lymphocytes. *J. Exp. Med.* **191**, 1881–1894 (2000).
28. Stephenson, I. *et al.* Detection of anti-H5 responses in human sera by HI using horse erythrocytes following MF59-adjuvanted influenza A/Duck/Singapore/97 vaccine. *Virus Res.* **103**, 91–95 (2004).

LETTERS

CLEC5A is critical for dengue-virus-induced lethal disease

Szu-Ting Chen¹, Yi-Ling Lin^{2,3}, Ming-Ting Huang¹, Ming-Fang Wu¹, Shih-Chin Cheng¹, Huan-Yao Lei⁴, Chien-Kuo Lee⁵, Tzyy-Wen Chiou⁶, Chi-Huey Wong³ & Shie-Liang Hsieh^{1,3,7}

Dengue haemorrhagic fever and dengue shock syndrome, the most severe responses to dengue virus (DV) infection, are characterized by plasma leakage (due to increased vascular permeability) and low platelet counts^{1,2}. CLEC5A (C-type lectin domain family 5, member A; also known as myeloid DAP12-associating lectin (MDL-1))³ contains a C-type lectin-like fold similar to the natural-killer T-cell C-type lectin domains and associates with a 12-kDa DNAX-activating protein (DAP12)⁴ on myeloid cells. Here we show that CLEC5A interacts with the dengue virion directly and thereby brings about DAP12 phosphorylation. The CLEC5A–DV interaction does not result in viral entry but stimulates the release of proinflammatory cytokines. Blockade of CLEC5A–DV interaction suppresses the secretion of proinflammatory cytokines without affecting the release of interferon- α , supporting the notion that CLEC5A acts as a signalling receptor for proinflammatory cytokine release. Moreover, anti-CLEC5A monoclonal antibodies inhibit DV-induced plasma leakage, as well as subcutaneous and vital-organ haemorrhaging, and reduce the mortality of DV infection by about 50% in STAT1-deficient mice. Our observation that blockade of CLEC5A-mediated signalling attenuates the production of proinflammatory cytokines by macrophages infected with DV (either alone or complexed with an enhancing antibody) offers a promising strategy for alleviating tissue damage and increasing the survival of patients suffering from dengue haemorrhagic fever and dengue shock syndrome, and possibly even other virus-induced inflammatory diseases.

Dengue is a mosquito-borne infection caused by four serotypes of dengue virus (DV1–DV4) and is currently the most common arboviral disease worldwide^{1,2}. Primary infection with any of the four DV serotypes typically results in mild dengue fever (DF) and provides lifelong immunity to the infecting strain. However, secondary infection with different DV serotypes is associated with an increased risk of developing dengue haemorrhagic fever (characterized by thrombocytopenia and capillary leakage) and can progress to life-threatening hypovolaemic dengue shock syndrome¹. The pathogenesis of dengue haemorrhagic fever and dengue shock syndrome remains unclear, but massive cytokine secretion (cytokine storm) is believed to be one of the major contributory factors⁵. Indeed, dengue is a major public health problem, with about 50 million people infected each year (of whom about 20,000 die) and about 2.5 billion people worldwide being at risk of infection. Unfortunately, no DV-specific therapies or vaccines are available^{1,2}.

Dendritic cells (DCs) and macrophages are the primary targets of DV infections^{6–8}. Whereas infected DCs undergo apoptosis (despite the secretion of proinflammatory cytokines by bystander DCs)⁹, infected macrophages survive for at least 45 days and secrete multiple

cytokines and chemokines from 6 h after infection¹⁰. This result suggests that macrophages are the major source of proinflammatory cytokines after infection with DV, in which virions might trigger inflammatory reactions by activating pattern recognition receptors. Toll-like receptors (TLRs), C-type lectins and immunoglobulin-like (Ig-like) receptors (for example, TREMs (triggering receptors expressed on myeloid cells) and TREM-like transcript (TLT)) have been implicated as potential pattern recognition receptors^{11–13}. To determine whether DV binds to and activates candidate pattern recognition receptors on immune cells, we expressed 22 fusion proteins in mammalian cells and screened for their interaction with DV2 (Supplementary Table 1). Among the receptors tested, the C-type lectin receptor DC-SIGN (dendritic-cell-specific intercellular adhesion molecule-3-grabbing non-integrin; also known as CLEC4L) has previously been shown to interact with glycans located on the envelope (E) protein of DV¹⁴. Using an enzyme-linked immunosorbent assay (ELISA) we showed that CLEC5A.Fc (in addition to DC-SIGN.Fc and DC-SIGNR.Fc) was able to capture DV2 (Fig. 1a). To confirm the specificity of the interaction between CLEC5A and DV, complexes were immunoprecipitated with Protein A–Sepharose beads and then probed with an anti-DV envelope (anti-E) monoclonal antibody (mAb). E protein was detected in the immunoprecipitates of DC-SIGN.Fc and CLEC5A.Fc, confirming that CLEC5A interacts with the dengue virion (Fig. 1b). However, whereas the binding of DC-SIGN to DV is Ca²⁺ dependent, EDTA (a Ca²⁺ chelator) had no effect on the CLEC5A–DV interaction (Fig. 1c). The Ca²⁺-independent binding feature is in accord with that of β -glucan receptor dectin 1, which also contains a similar C-type lectin-like domain and binds zymosan independently of metal ions¹⁵. Furthermore, transfection of 293T cells with DC-SIGN and CLEC5A resulted in increased binding of biotinylated DV to the cells (Fig. 1d). There are two conserved N-linked glycosylation sites at Asn 67 and Asn 153 of E protein¹⁴, and these N-glycans have been implicated in cellular attachment and viral entry¹⁶. To investigate the participation of glycans in the association of CLEC5A with DV, virions were incubated with fucose, mannose or mannan (the latter two sugars are ligands for DC-SIGN¹⁷). As expected, mannose and mannan caused dose-dependent inhibition of the interaction between DC-SIGN and DV (Fig. 1e). The binding of CLEC5A to DV was significantly reduced in the presence of fucose ($P < 0.0001$) and, to a smaller extent, mannan ($P = 0.0005$) (Fig. 1e), indicating that fucose might be more important than mannose in the CLEC5A–DV interaction.

DC-SIGN, which is expressed on DCs and macrophages (Supplementary Fig. 1a), contains three motifs in its cytoplasmic tail that are believed to be involved in endocytosis or intracellular

¹Department and Institute of Microbiology and Immunology, National Yang-Ming University, Taipei 112, Taiwan. ²Institute of Biomedical Sciences, and ³Genomics Research Center, Academia Sinica, Taipei 115, Taiwan. ⁴Department of Microbiology and Immunology, National Cheng Kung University, Tainan 701, Taiwan. ⁵Institute of Immunology, National Taiwan University, Taipei 100, Taiwan. ⁶Department of Life Science and Institute of Biotechnology, National Dong Hwa University, Hualien 974, Taiwan. ⁷Immunology Research Center, National Yang-Ming University & Taipei Veterans General Hospital, Taipei 112, Taiwan.

trafficking¹⁸. In contrast, CLEC5A was originally identified as a DAP12-associated molecule expressed exclusively on monocytes and macrophages (Supplementary Fig. 1b), although its ligand(s) and biological functions remain to be determined³. Infection of macrophages with DV was found to induce DAP12 phosphorylation in a dose-dependent manner (Fig. 2a). DAP12 phosphorylation peaked at 12 h after infection and persisted for at least 48 h in the presence of live DV, whereas ultraviolet-inactivated DV (UV-DV) triggered only limited DAP12 phosphorylation that lasted for 12 h (Fig. 2b). This indicates that DAP12 phosphorylation is independent of DV replication during the first 2–6 h of infection. Knockdown of CLEC5A (using the short hairpin RNA (shRNA) pLL3.7/CLEC5A), but not that of DC-SIGN (by pLL3.7/DC-SIGN), caused a substantial decrease in DAP12 phosphorylation (Fig. 2c), suggesting that DV-triggered DAP12 phosphorylation is mediated through CLEC5A. It is known that DC-SIGN participates in the infection of DCs by DV¹⁹.

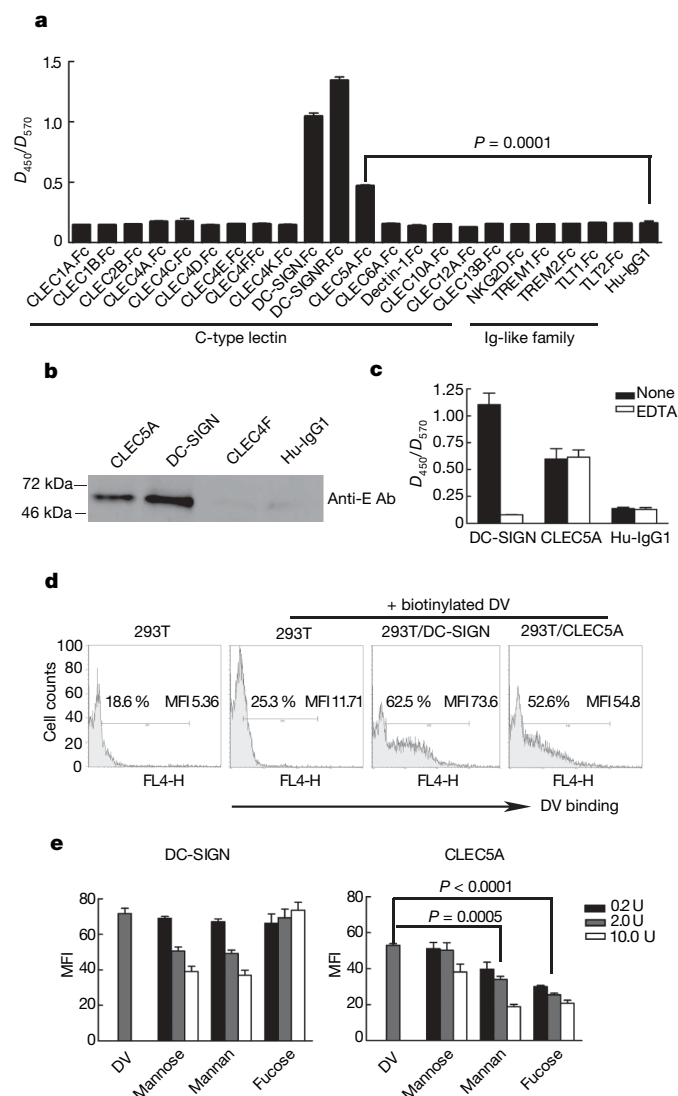


Figure 1 | DV interacts with CLEC5A. **a**, Interaction of DV with receptor.Fc was determined by ELISA. **b**, DV–receptor.Fc complexes were immunoprecipitated and detected by western blotting. **c**, Inhibition of CLEC5A–DV interaction by EDTA was determined by ELISA. **d**, Both DC-SIGN and CLEC5A increase DV binding to 293T cells. **e**, Addition of sugars inhibits binding of biotinylated DV to DC-SIGN-transfected or CLEC5A-transfected 293T cells. MFI, mean fluorescence intensity. Units (U) for monosaccharide (mannose and fucose) and polysaccharide (mannan) are mM and mg ml⁻¹, respectively. Data are expressed as means and s.d. for three independent experiments. Two-tailed Student's *t*-tests were performed.

We therefore tested the hypothesis that CLEC5A is involved in the entry of DV into macrophages by monitoring the expression of DV nonstructural protein 3 (NS3), which is expressed when DV replicates in macrophages. In contrast to that of DC-SIGN, knockdown of CLEC5A by shRNA (Fig. 2d) or blocking of the CLEC5A–DV interaction with anti-CLEC5A antibody (Fig. 2e) did not inhibit NS3 expression in macrophages as examined by flow cytometry and confocal microscopy, respectively. The shRNA pLL3.7/CLEC5A also failed to suppress the release of dengue virions into the supernatant of infected macrophages, as determined by a plaque-forming assay (data not shown). These results indicate that, whereas DC-SIGN mediates DV infection and replication, the interaction of DV with CLEC5A triggers cell signalling.

To determine whether CLEC5A is involved in DV-induced inflammation, we examined the secretion of inflammatory cytokines by macrophages after infection with DV. At 6 h after infection, we detected dose-dependent secretion of tumour necrosis factor (TNF)- α ; similar levels of cytokine were secreted by macrophages infected with either DV or UV-DV (Fig. 3a, left panel). However,

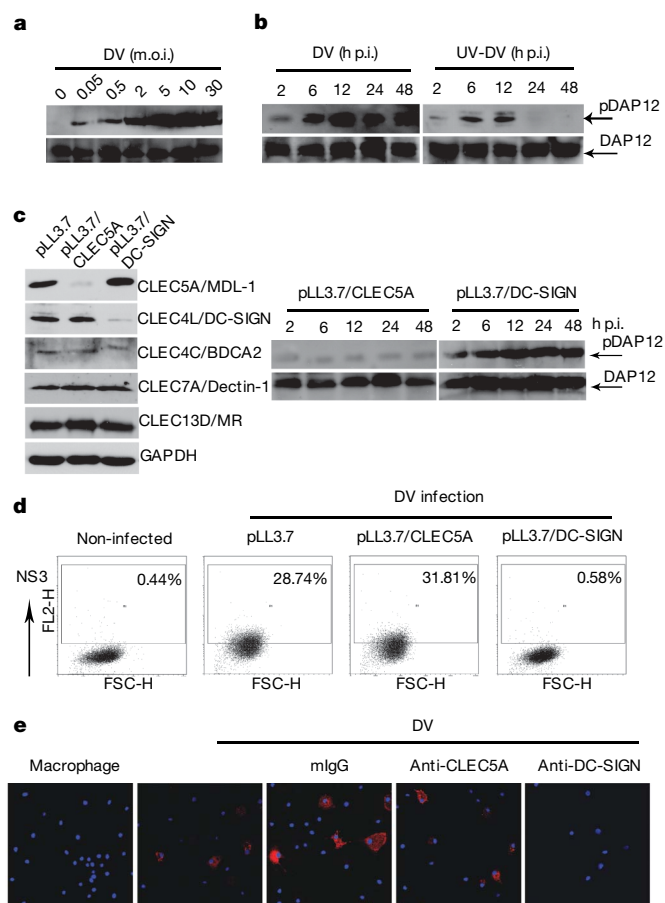


Figure 2 | CLEC5A is essential for DV-induced DAP12 phosphorylation, but not for DV replication. **a**, DV-induced DAP12 phosphorylation (2 h after infection) in human macrophages was determined by western blotting. m.o.i., multiplicity of infection. **b**, Kinetics of DAP12 phosphorylation induced by DV and ultraviolet-inactivated DV (UV-DV) was determined by western blotting. p.i., after infection. **c**, **d**, Effects of shRNAs on protein expression and inhibition of DV-mediated DAP12 phosphorylation (**c**) and DV entry and replication in macrophages (**d**). GAPDH, glyceraldehyde-3-phosphate dehydrogenase. **e**, The effects of antagonistic mAbs in inhibiting NS3 expression (red) in DV-infected macrophage at 48 h after infection. Cells were counterstained with Hoechst 33342 (blue) and were examined with a confocal microscope (m.o.i. = 5). The second panel shows the cells (macrophages) infected with dengue virus (DV) only, without any other treatment.

at 12 h after infection, TNF- α secretion was further increased by DV but not by UV-DV (Fig. 3a, right panel). Over a 48-h time course, TNF- α secretion continually increased for macrophages infected with DV, whereas at 24–48 h after infection with UV-DV this cytokine was barely detectable (Fig. 3b). These data correlate with the kinetics of DAP12 phosphorylation (Fig. 2b), indicating that DV-mediated secretion of TNF- α is related to DAP12 activation. We also observed that knockdown of CLEC5A suppressed the release of TNF- α , interleukin (IL)-6, IL-8, macrophage inflammatory protein (MIP)-1 α and interferon-inducible protein (IP)-10 by DV-infected macrophages to a much greater extent than knockdown of DC-SIGN (Fig. 3c and Supplementary Fig. 2). However, whereas pLL3.7/DC-SIGN mildly suppressed interferon (IFN)- α secretion ($P = 0.048$), pLL3.7/CLEC5A did not affect this cytokine (Fig. 3c). To further dissect the DV-activated signalling pathways leading to cytokine secretion, macrophages were transfected with shRNAs to knock down CLEC5A, DC-SIGN, TLR4, TLR7 or MyD88, before DV infection. The data indicate that DV-induced IFN- α secretion occurs through the TLR7–MyD88 pathway ($P = 0.0016$), whereas TNF- α secretion is mediated through both CLEC5A ($P = 0.003$) and TLR7–MyD88 ($P = 0.013$) (Fig. 3d). We generated a panel of

anti-CLEC5A mAbs with differential antagonistic effects on the four serotypes of DV (DV1–DV4; Supplementary Table 2), which were determined by inhibition of TNF- α secretion from DV-infected macrophages (Fig. 3e). Although different epitopes of CLEC5A may mediate their individual interactions, all antibodies that inhibit the CLEC5A–DV interaction suppress the inflammatory response by macrophages infected with relevant DV serotype(s). The differential antagonistic effects of anti-CLEC5A mAbs might be related to heterogeneous glycosylation among DV serotypes²⁰. Additionally, an anti-CLEC5A mAb may inhibit the binding of a specific DV serotype whose binding site overlaps with that of anti-CLEC5A mAb.

It has previously been shown that non-neutralizing anti-DV Abs promote DV entry into target cells by means of FcR receptors and thereby enhance cytokine release^{6,21}, a phenomenon termed antibody-dependent enhancement (ADE) of infection. For example, anti-prM and anti-E mAbs have been shown to induce this effect *in vitro*²². Here we investigated whether blockade of the CLEC5A–DV interaction can inhibit ADE. Primary human macrophages were infected with DV alone or with anti-prM–DV or anti-E–DV immunocomplexes, in the presence of anti-CLEC5A mAb (or isotype control) for 36 h. Anti-prM–DV and anti-E–DV immunocomplexes (ADE)

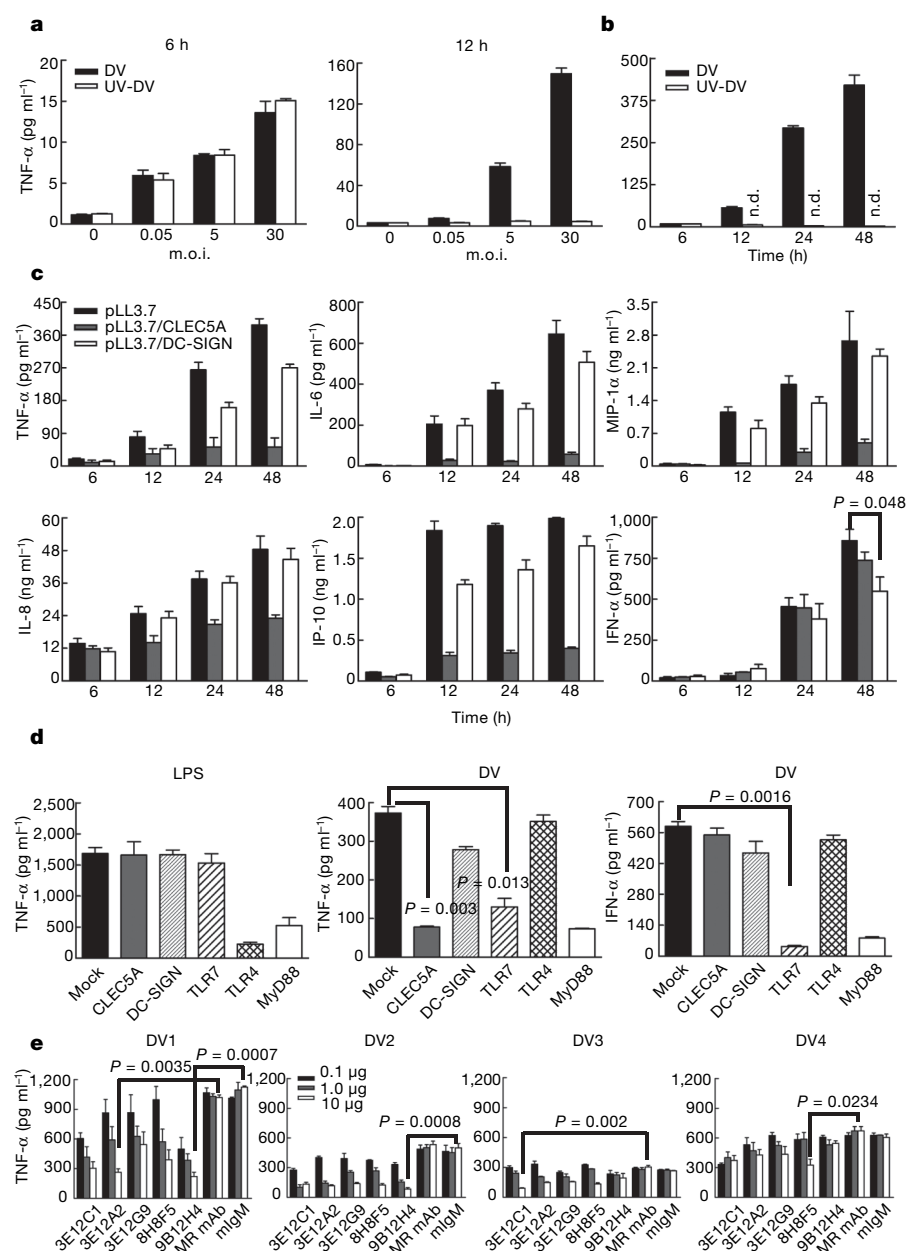


Figure 3 | CLEC5A is critical for DV-mediated secretion of TNF- α but not that of IFN- α . **a**, The dose dependence of DV-induced and UV-DV-induced TNF- α secretion by macrophages was measured by ELISA at 6 and 12 h after infection. **b**, Kinetics of TNF- α expression after DV infection. n.d., not detectable. **c**, Effects of CLEC5A and DC-SIGN shRNAs on the secretion of cytokines from DV-infected macrophages. **d**, Knockdown experiments with specific shRNAs show that DV-induced IFN- α secretion is via the TLR7–MyD88 pathway. **e**, TNF- α secretion in response to DV serotypes 1–4 is inhibited by antagonistic anti-CLEC5A mAbs (Supplementary Table 2), M.R. (mannose receptor); mIgM (murine IgM isotype control). (m.o.i. = 5). Data are expressed as means and s.d. for three independent experiments.

increased the expression of NS3 (Supplementary Fig. 3a) and the levels of TNF- α and IFN- α secretion in comparison with DV alone (Supplementary Fig. 3b, c). However, whereas anti-CLEC5A mAb significantly inhibited TNF- α release from macrophages infected with DV, or from macrophages incubated with anti-prM-DV or anti-E-DV immunocomplexes (Supplementary Fig. 3c), the secretion of IFN- α was not affected, suggesting that ADE-mediated IFN- α secretion is independent of CLEC5A. This result supports the notion that CLEC5A acts in a signalling capacity but does not alter relative infection levels.

We found that inhibition of the CLEC5A–DV interaction attenuated DV-induced permeability changes (Supplementary Fig. 4) *in vitro*. We therefore further investigated whether blockade of the CLEC5A–DV interaction can rescue mice from DV-induced lethality *in vivo*. We found that murine CLEC5A (mCLEC5A) binds to DV with an affinity similar to that of human CLEC5A (Supplementary Fig. 5a) and is expressed on myeloid lineages (CD11b⁺, F4/80⁺), bone marrow-derived macrophages and murine macrophage-like Raw264.7 cells (Supplementary Fig. 5b, c). Furthermore, DV replicated more efficiently and stimulated greater TNF- α secretion in bone marrow-derived macrophages derived from STAT1-deficient (STAT1^{-/-})²³ mice than in wild-type mice (Supplementary Fig. 6a, b), and blockade of mCLEC5A–DV interaction by antagonistic mAbs (Supplementary Table 3) abolished DV-induced TNF- α secretion by bone marrow-derived macrophages in a dose-dependent manner (Supplementary Fig. 6c). This indicates that the function of mCLEC5A is similar to that of human CLEC5A.

IFN- α functions to inhibit viral replication in both infected and uninfected cells, and IFN-mediated responses to DV infection

involve both the STAT1-dependent (essential in the control of viral replication) and STAT1-independent (essential for the resolution of infection) pathways²⁴. Although wild-type mice were resistant to infection with DV, STAT1^{-/-} mice were sensitive to lethality induced by DV2 (strain New Guinea C-N; Supplementary Fig. 7). We tested the potential therapeutic effects of the antagonistic mAbs on STAT1^{-/-} mice. DV-challenged STAT1^{-/-} mice showed ruffled fur and mild paralysis in addition to subcutaneous and intestinal haemorrhaging at 8 days after infection (Fig. 4a), and they all died within 7–14 days of infection (Fig. 4e). Five doses of antibodies (100 μ g per mouse, intraperitoneally) or TNFR2.Fc (100 μ g per mouse, intraperitoneally) were administered on days 0, 1, 3, 5 and 7 after infection. At 9 days after infection, leakage of Evans blue into the kidney, liver, stomach, small intestine and spleen of DV-challenged mice was significantly decreased in mice treated with anti-mCLEC5A mAbs in comparison with controls (Fig. 4b, c). Anti-mCLEC5A mAbs also effectively lowered the serum levels of TNF- α and IP-10 (Fig. 4d, top and middle panels), without suppressing viral replication, at day 7 after infection (Fig. 4d, bottom panel), and protected mice from lethality at day 14 after infection (70% protection rate). The overall survival rate of anti-mCLEC5A-treated mice was 48% at day 21 after infection (Fig. 4e), with DV being cleared from serum of the surviving mice at day 23 after infection (data not shown). Thus, blocking the CLEC5A–DV interaction seems to prevent the DV-associated complications of haemorrhaging and plasma leakage. It also suppresses the macrophage inflammatory response without impairing virus clearance by the adaptive immune response. In contrast, TNFR2.Fc neither reduced

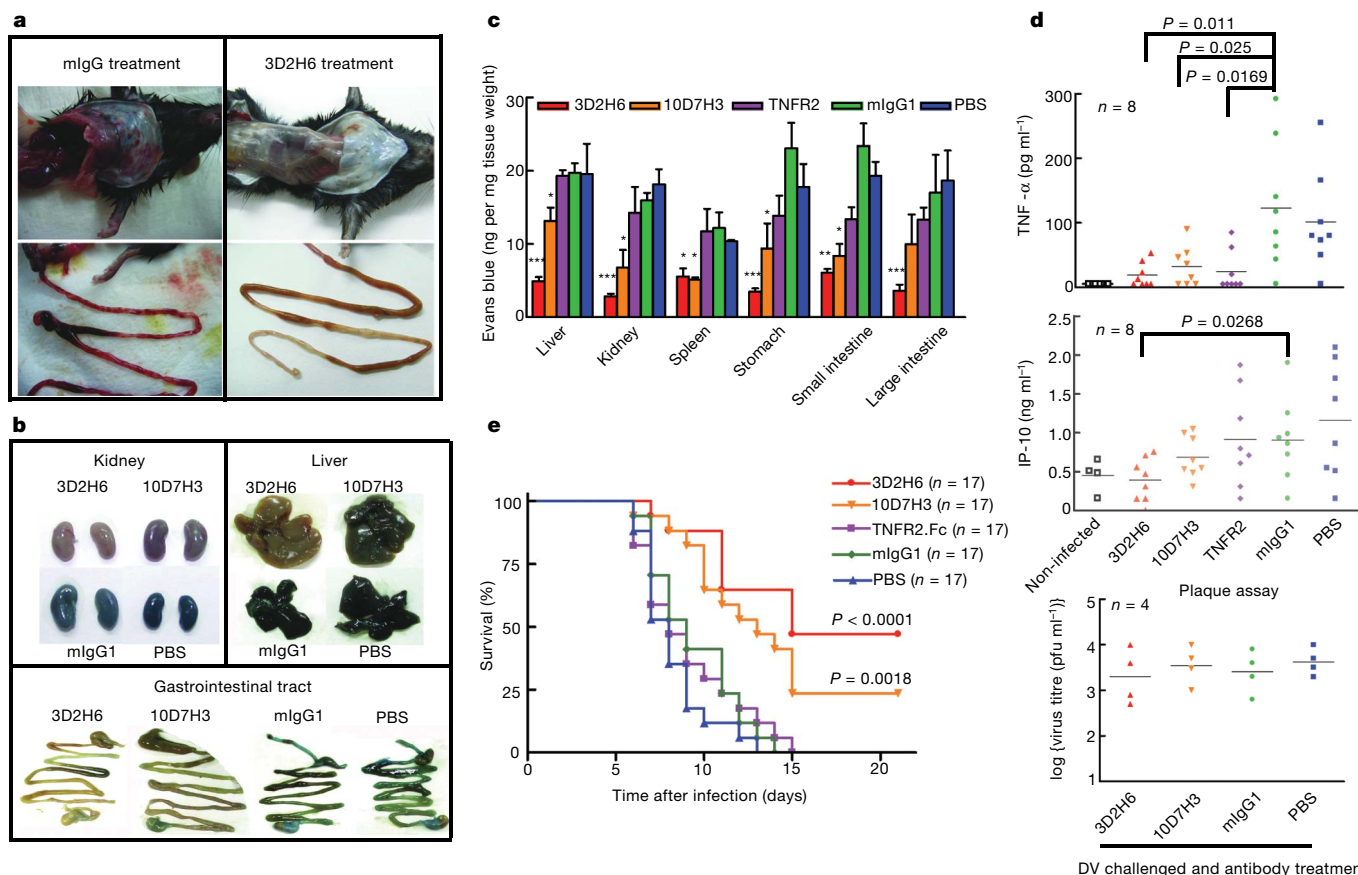


Figure 4 | Anti-mCLEC5A mAbs prevents DV-induced vascular leakage and lethality in STAT1-deficient mice. **a**, **b**, Anti-murine CLEC5A mAb inhibits subcutaneous and intestinal haemorrhaging (**a**) and plasma leakage (**b**) of DV-challenged mice. **c**, Quantification of Evans blue extracted from organs. Data are expressed as means and s.d. for three independent experiments. Asterisk, $P < 0.05$; two asterisks, $P < 0.01$; three asterisks,

$P < 0.001$ (Student's *t*-test). **d**, Serum levels of cytokines and virus titres at day 7 after infection (two-tailed Student's *t*-tests). p.f.u., plaque-forming units. **e**, Survival curve of mice challenged with DV2 (New Guinea C-N). Data were collected from four independent experiments and are shown as Kaplan–Meier survival curves with log rank test.

vascular permeability (Fig. 4c) nor protected mice from lethality (Fig. 4e), despite effectively lowering the level of TNF- α in serum (Fig. 4d).

We have thus demonstrated that CLEC5A is an important pattern recognition receptor in the context of dengue virus infection. It is clear that although DV-triggered cytokine release from macrophages involves both CLEC5A and TLR7 pathways, only the blockade of CLEC5A–DV interactions can attenuate inflammation and maintain host immunity so as to clear virus. Even though the direct relevance of this *STAT1*^{−/−} mouse model in relation to human infection with DV remains unclear, our *in vivo* protection data of CLEC5A mAbs still shed light on its therapeutic potential for blocking the progression of severe dengue disease in patients with dengue haemorrhagic fever or dengue shock syndrome. It will be interesting to test whether this strategy can be applied to the identification of receptors interacting with other pathogenic viruses, and to the treatment of the severe inflammatory consequences of infection.

METHODS SUMMARY

Reagents. Human MyD88, TLR4 and TLR7 shRNA expression vectors were from InvivoGen. The DC-SIGN shRNA expression vector was provided by P. Vincet²⁵ and TNFR2.Fc (Enbrel; Wyeth) was provided by H.-Y. Lin. Anti-CLEC4L/DC-SIGN mAb (clone 120507), anti-CLEC4C/BDCA2 and anti-CLEC7A/dectin-1 polyclonal antibody were purchased from R&D Systems. Anti-human CD206 (mannose receptor; clone 15-2; IgG1) was purchased from BioLegend.

Virus stock. Four DV serotypes, namely DV1/Hawaii, DV2/PL046, DV3/H-087 and DV4/866146A (ref. 26), were used for *in vitro* studies. For animal studies, mouse-adapted neurovirulent DV2 (strain New Guinea C-N) was provided by C. J. Lai. Virus propagation was performed in C6/36 cells, and virus titres were determined by plaque-forming assays with BHK-21 cells. Unless otherwise specified, DV2/PL046 was used throughout the study.

Transfection of macrophages. Macrophages were transfected by electroporation with the Amaxa Human Monocyte Nucleofector kit (Amaxa Inc.) in accordance with the manufacturer's instructions. In brief, macrophages (6×10^6) were harvested and resuspended in 100 μ l of nucleofector solution. After addition of appropriate shRNA expression vector or control vector (5 μ g), cells were electroporated with Amaxa program Y-001 and recovered for 16 h before infection with DV. The transfection efficiency was determined by fluorescence microscopy, and the transfected (enhanced green fluorescent protein-positive) cells were sorted by FACSaria (Becton Dickinson) before cytokine assay.

Permeability assay. Vascular leakage was examined by intravascular administration of Evans blue into mice, as described previously²⁷. Uptake of Evans blue by organs was quantified by extracting with formamide and measuring the attenuation at 610 nm (D_{610}).

Inoculation of virus. Mice were challenged intraperitoneally with 10^5 plaque-forming units of DV2 (New Guinea C-N) in 300 μ l of PBS and simultaneously injected intracranially (i.c.) with 30 μ l of PBS into the right hemisphere of mouse brains.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 29 February; accepted 18 April 2008.

Published online 21 May 2008.

- Wilder-Smith, A. & Schwartz, E. Dengue in travelers. *N. Engl. J. Med.* **353**, 924–932 (2005).
- Mackenzie, J. S., Gubler, D. J. & Petersen, L. R. Emerging flaviviruses: the spread and resurgence of Japanese encephalitis, West Nile and dengue viruses. *Nature Med.* **10**, S98–S109 (2004).
- Bakker, A. B., Baker, E., Sutherland, G. R., Phillips, J. H. & Lanier, L. L. Myeloid DAP12-associating lectin (MDL)-1 is a cell surface receptor involved in the activation of myeloid cells. *Proc. Natl Acad. Sci. USA* **96**, 9792–9796 (1999).
- Lanier, L. L., Corliss, B. C., Wu, J., Leong, C. & Phillips, J. H. Immunoreceptor DAP12 bearing a tyrosine-based activation motif is involved in activating NK cells. *Nature* **391**, 703–707 (1998).
- Pang, T., Cardoso, M. J. & Guzman, M. G. Of cascades and perfect storms: the immunopathogenesis of dengue haemorrhagic fever–dengue shock syndrome (DHF/DSS). *Immunol. Cell Biol.* **85**, 43–45 (2007).

- Halstead, S. B. & O'Rourke, E. J. Dengue viruses and mononuclear phagocytes. I. Infection enhancement by non-neutralizing antibody. *J. Exp. Med.* **146**, 201–217 (1977).
- Palucka, A. K. Dengue virus and dendritic cells. *Nature Med.* **6**, 748–749 (2000).
- Wu, S. J. *et al.* Human skin Langerhans cells are targets of dengue virus infection. *Nature Med.* **6**, 816–820 (2000).
- Palmer, D. R. *et al.* Differential effects of dengue virus on infected and bystander dendritic cells. *J. Virol.* **79**, 2432–2439 (2005).
- Chen, Y. C. & Wang, S. Y. Activation of terminally differentiated human monocytes/macrophages by dengue virus: productive infection, hierarchical production of innate cytokines and chemokines, and the synergistic effect of lipopolysaccharide. *J. Virol.* **76**, 9877–9887 (2002).
- Cook, D. N., Pisetsky, D. S. & Schwartz, D. A. Toll-like receptors in the pathogenesis of human disease. *Nature Immunol.* **5**, 975–979 (2004).
- Klesney-Tait, J., Turnbull, I. R. & Colonna, M. The TREM receptor family and signal integration. *Nature Immunol.* **7**, 1266–1273 (2006).
- Robinson, M. J., Sancho, D., Slack, E. C., LeibundGut-Landmann, S. & Reis e Sousa, C. Myeloid C-type lectins in innate immunity. *Nature Immunol.* **7**, 1258–1265 (2006).
- Pokidysheva, E. *et al.* Cryo-EM reconstruction of dengue virus in complex with the carbohydrate recognition domain of DC-SIGN. *Cell* **124**, 485–493 (2006).
- Brown, G. D. & Gordon, S. A new receptor for β -glucans. *Nature* **413**, 36–37 (2001).
- Modis, Y., Ogata, S., Clements, D. & Harrison, S. C. Variable surface epitopes in the crystal structure of dengue virus type 3 envelope glycoprotein. *J. Virol.* **79**, 1223–1231 (2005).
- Mitchell, D. A., Fadden, A. J. & Drickamer, K. A novel mechanism of carbohydrate recognition by the C-type lectins DC-SIGN and DC-SIGNR. Subunit organization and binding to multivalent ligands. *J. Biol. Chem.* **276**, 28939–28945 (2001).
- Lozach, P. Y. *et al.* Dendritic cell-specific intercellular adhesion molecule 3-grabbing non-integrin (DC-SIGN)-mediated enhancement of dengue virus infection is independent of DC-SIGN internalization signals. *J. Biol. Chem.* **280**, 23698–23708 (2005).
- Tassaneeritthep, B. *et al.* DC-SIGN (CD209) mediates dengue virus infection of human dendritic cells. *J. Exp. Med.* **197**, 823–829 (2003).
- Johnson, A. J., Guirakhoo, F. & Roehrig, J. T. The envelope glycoproteins of dengue 1 and dengue 2 viruses grown in mosquito cells differ in their utilization of potential glycosylation sites. *Virology* **203**, 241–249 (1994).
- Gonzalez, A. P., Engle, R. E., St Claire, M., Purcell, R. H. & Lai, C. J. Monoclonal antibody-mediated enhancement of dengue virus infection *in vitro* and *in vivo* and strategies for prevention. *Proc. Natl Acad. Sci. USA* **104**, 9422–9427 (2007).
- Huang, K. J. *et al.* The dual-specific binding of dengue virus and target cells for the antibody-dependent enhancement of dengue virus infection. *J. Immunol.* **176**, 2825–2832 (2006).
- Durbin, J. E., Hackenmiller, R., Simon, M. C. & Levy, D. E. Targeted disruption of the mouse Stat1 gene results in compromised innate immunity to viral disease. *Cell* **84**, 443–450 (1996).
- Shresta, S. *et al.* Critical roles for both STAT1-dependent and STAT1-independent pathways in the control of primary dengue virus infection in mice. *J. Immunol.* **175**, 3946–3954 (2005).
- Arrighi, J. F. *et al.* Lentivirus-mediated RNA interference of DC-SIGN expression inhibits human immunodeficiency virus transmission from dendritic cells to T cells. *J. Virol.* **78**, 10848–10855 (2004).
- Lin, Y. L. *et al.* Study of Dengue virus infection in SCID mice engrafted with human 562 cells. *J. Virol.* **72**, 9729–9737 (1998).
- Shresta, S., Sharar, K. L., Prigozhin, D. M., Beatty, P. R. & Harris, E. Murine model for dengue virus-induced lethal disease with increased vascular permeability. *J. Virol.* **80**, 10208–10217 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank W.-C. Yeh, C. Milner and J. Paulson for critical comments; N.-J. Chen, C.-H. Lin, Y.-L. Lee and J.-J. Liang for technical assistance. Resources and collaborative efforts were provided by the RNAi Consortium, Academia Sinica, Taiwan, and the Consortium for Functional Glycomics funded by the National Institute of General Medical Sciences (GM62116). This work was supported mainly by the National Research Program for Genomic Medicine, National Science Council, Taiwan (NSC-95-3112-B-010-0171 and NSC 96-3112-B-010-2), and in part by the National Yang-Ming University, Taiwan (96A-D-D132 from the Ministry of Education), Taipei Veterans General Hospital (V97S5-001), and Academia Sinica.

Author Contributions S.-T.C. designed, performed and analysed experiments, and wrote the paper. Y.-L.L. designed, analysed experiments and wrote the paper. M.-T.H., M.-F.W. and S.-C.C. performed experiments. H.-Y.L., C.-K.L. and T.-W.C. provided materials and reagents. C.-H.W. analysed experiments. S.-L.H. designed and analysed experiments, and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.-L.H. (slhsieh@ym.edu.tw or slhsieh@gate.sinica.edu.tw).

METHODS

Reagents. Evans blue and human IgG1 mAbs were purchased from Sigma. Human macrophage-colony-stimulating factor (M-CSF) and TNF- α were purchased from R&D Systems. Protein A–Sepharose beads, horseradish peroxidase (HRP)-conjugated second antibody and enhanced chemiluminescence reagents were from Amersham Biosciences. Proteinase inhibitor cocktail tablets were from Roche. Anti-CD14 microbeads were from Miltenyi Biotec GmbH. Anti-phosphotyrosine mAb (clone 4G10) was purchased from Upstate Biotechnology. The polyclonal antibody against human DAP12 was from Santa Cruz Biotechnology.

Biotinylation of dengue virions. Sucrose-cushion-purified DV2 (about 5×10^9 plaque-forming units) were resuspended in PBS, followed by the addition of 80 μ l of 10 mM biotin solution (EZ-Link Sulfo-NHS-SS-Biotin; Cat. 21331; Pierce) and incubated on ice for 2 h. After removal of the unconjugated biotin by dialysis, the biotinylated virions were stored at 4 °C, and the efficiency of biotinylation was determined by flow cytometry with allophycocyanin-conjugated streptavidin.

Receptor-virus interaction. Receptor.Fc constructs in pcDNA3.1 (Invitrogen) were transfected into 293 FreeStyle cells (Invitrogen) and expressed proteins were purified with Protein A beads. Fusion proteins (1 μ g per well) were coated on microtitre plates to capture DV2 (5×10^6 plaque-forming units per well), and the bound dengue virions were detected by biotinylated anti-DV envelope protein antibody²⁶ and HRP-conjugated streptavidin with 3,3',5,5'-tetramethylbenzidine (TMB) (BD Pharmingen) substrate.

Sugar competition assay. Human 293T cells overexpressing human DC-SIGN or CLEC5A were preincubated with 15 U of heparin for 15 min, followed by the addition of biotinylated dengue virus (m.o.i. = 20) in conjunction with monosaccharides (mannose and fucose) or polysaccharide (mannan) and incubated at 4 °C for 30 min. After being washed with PBS, cells were incubated with allophycocyanin-conjugated streptavidin at 4 °C for 30 min, and the biotinylated dengue virus adsorbed on cell surfaces was detected by flow cytometry.

Preparation of PBMC and *in vitro* differentiation of monocytes into macrophage and DCs. Peripheral blood mononuclear cells (PBMCs) were isolated from the whole blood of healthy human donors by standard density-gradient centrifugation with Ficoll-Paque (Amersham Biosciences). Purified neutrophils were resuspended in PBS (pH 7.4) and hypotonic lysis of erythrocytes was performed to improve the purity of polymorphonuclear cells. CD14⁺ cells were purified from PBMCs by high-gradient magnetic sorting, using the VarioMACS technique with anti-CD14 microbeads (Miltenyi Biotec GmbH). Cells were then cultured in complete RPMI 1640 medium (Life Technologies) supplemented with 10 ng ml⁻¹ human M-CSF (R&D Systems) for 6 days (ref. 28). DCs were generated from adherent PBMCs by incubation in RPMI 1640 medium supplemented with 10% FCS, 800 U ml⁻¹ human GM-CSF (Leucomax; Schering-Plough) and 500 U ml⁻¹ human IL-4 (R&D Systems) for 6 days (immature DCs). To prepare mature activated DCs, immature DCs were further incubated for 36 h with γ -irradiated (5,500 rad) CD40 ligand (CD40L)-expressing L cells (DNAX Research Institute) at a ratio of 3:1 (ref. 29).

Flow cytometry analysis. Human polymorphonuclear cells and PBMCs were isolated from the whole blood of healthy donors; macrophages and DCs were generated from CD14⁺ monocytes as described above. To characterize the expression pattern of CLEC5A and DC-SIGN, cells were stained with fluorescein isothiocyanate (FITC)-conjugated anti-CLEC5A mAb (clone 283834; R&D Systems) or FITC-conjugated anti-DC-SIGN mAb (BD Pharmingen) in conjunction with phycoerythrin (PE)-conjugated mAbs against CD3, CD19, CD56, CD14 or CD66 (BD Pharmingen). Results were compared with those obtained using isotype-matched controls (IgG2b for anti-CLEC5A mAb, IgG1 for anti-DC-SIGN; Sigma). To detect intracellular DV antigens, infected cells were fixed with 1% (v/v) paraformaldehyde and permeabilized with 0.1% (w/v) saponin, followed by staining with NS3 mAb or an isotype-matched control (mIgG1; Sigma). After incubation for 1 h, PE-conjugated goat F(ab')₂ anti-mouse IgG secondary antibody was added to detect emitted fluorescence by FACSCalibur (Becton Dickinson) with CellQuest software (Becton Dickinson).

Detection of NS3 in DV-infected cells by immunofluorescence assay. To determine the blocking effects of mAbs on DV entry to macrophages, cells were cultured on glass coverslips and preincubated with anti-DC-SIGN mAb (50 μ g ml⁻¹; clone 120507; R&D Systems), anti-CLEC5A mAb (50 μ g ml⁻¹; clone 8H8F5), or mouse IgG1 (50 μ g ml⁻¹; Sigma-Aldrich) for 2 h at 4 °C, before the addition of DV2. After incubation with DV2 for 2.5 h, and unbound virus was removed by washing the cell monolayers with serum-free RPMI and then with complete RPMI medium containing blocking antibodies. After washing, cells were fixed with 4% (v/v) paraformaldehyde, permeabilized with 0.5% (v/v) Triton X-100 in PBS for 10 min, and then incubated with blocking buffer [10% (w/v) BSA in PBS], before the addition of anti-NS3 mAb (20 μ g ml⁻¹). After

being washed, cells were incubated with Cy3-conjugated donkey anti-mouse IgG (Jackson Immuno) and Hoechst 33342 to detect NS3 and nuclei, respectively. Slides were mounted and observed by fluorescence microscopy (AX-70 laser scanning microscope; Olympus).

Detection of DAP12 by immunoprecipitation and immunoblotting. Macrophages were stimulated with dengue virus, followed by resuspension in lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% (v/v) Triton X-100, 0.1% (w/v) SDS, 5 mM EDTA, 10 mM NaF, 1 mM sodium orthovanadate, and proteinase inhibitor cocktail tablet (Roche)). Equal amounts of total cell extracts (100 μ g) were incubated with rabbit anti-DAP12 polyclonal antibody (Santa Cruz Biotechnology Inc.) at 4 °C for 4 h and then with 15 μ l of Protein A–Sepharose (Amersham Biosciences AB) for 2 h. The immunocomplex was washed three times before fractionation on SDS-PAGE, followed by transfer to nitrocellulose membrane and probing with anti-phosphotyrosine antibody (4G10; Upstate Biotechnology, Inc.). Immunoblots were developed with HRP-conjugated anti-mouse IgG antiserum (Cat. AP181P; Chemicon) followed by enhanced chemiluminescence detection reagents (Amersham). To detect the total amount of DAP12 in the blot, the membrane was stripped with Re-Blot Plus Strong solution (Cat. 2504; Chemicon), and probed with rabbit anti-DAP12 antibody. The effects of pLL3.7/DC-SIGN and pLL3.7/CLEC5A at 24 h after transfection were determined by probing the blot with anti-CLEC5A and anti-DC-SIGN mAbs, respectively (R&D Systems).

RNA interference. The coding region of human CLEC5A was targeted with the sequence 5'-TTGTTGGAATGACCTTAT-3' (clone 5), which was adapted by addition of a loop sequence (5'-TTCAAGAGA-3') to create a shRNA. The polymerase III terminator stretch used here was 5'-TTTTTT-3'. The sequences of scrambled RNAs of clone 5, as well as other targeting sequences (clones 1, 2, 3, 4 and 6) is shown in Supplementary Fig. 2. The pLL3.7 gene silencing vector was provided by V. Parijs. For the human DC-SIGN shRNA, the backbone of the pSUPER-siDC-SIGN construct (provided by P. Vincet) was used to target the sequence 5'-AAGGCTGCAGTGGGTGAGCTT-3'. The shRNA was amplified by PCR, using sense and antisense primers tagged with *HpaI* and *XhoI* sites, respectively, and was subcloned into the pLL3.7 vector. Human MyD88 and TLR7 shRNA expression vectors were purchased from InVivoGen; the target sequences are 5'-AACTGGAACAGACAACTATC-3' (ref. 30) and 5'-GGCAGACCTTGGATCTAAGTA-3', respectively, where these are cloned in the psiRNA-h7SKGFPzeo backbone.

Generation and characterization of anti-CLEC5A mAbs and selection of anti-agonistic antibody. Breeder mice (BALB/c strain) were maintained in the standard animal facility of the National Yang-Ming University. For the production of mAbs, mice were immunized with purified recombinant CLEC5A.Fc fusion protein as antigen. The most suitable mouse was selected for administration of the final boost. Lymphocytes from the immunized mouse spleen were fused with mouse myeloma NS-1 cells in the presence of 50% (v/v) polyethylene glycol (PEG1450; Sigma). Fused cells were cultured in HAT selection medium and the medium was refreshed after one week. About 2 weeks after fusion, culture supernatants were screened by ELISA to identify the candidate clones for further analysis by limiting dilution. Anti-CLEC5A mAbs were selected by ELISA-based differential screening, and only those recognizing recombinant CLEC5A.Fc, but not human IgG1, were regarded as positive clones. A similar strategy was used to generate anti-mCLEC5A mAbs. To select antagonistic mAbs against human CLEC5A and murine CLEC5A, mAbs were incubated with human macrophages (6×10^4 per well) and Raw264.7 cells transfected with human DC-SIGN, respectively, in 96-well plates for 30 min at 37 °C, before the addition of DV2 (m.o.i. = 5) and incubation for 2.5 h. After washing out unbound virus, cells were cultured for a further 36 h before harvesting the supernatants to determine TNF- α levels by ELISA.

Preparation of murine bone marrow-derived macrophages. Bone marrow cells were isolated from femurs and tibias and were incubated in RPMI complete medium supplemented with 10% fetal calf serum (Gibco-BRL) and 10 ng ml⁻¹ recombinant mouse M-CSF (R&D) for 6–8 days. At day 7, the expression of F4/80 (murine macrophage marker) was examined by fluorescence-activated cell sorting. More than 90% of cells were F4/80⁺ under this culture condition.

Quantification of vascular permeability, serum TNF- α and virus titre in infected mice. Vascular leakage was examined by intravascular administration of Evans blue (Sigma-Aldrich) as described previously²⁷. In brief, Evans blue (0.15 ml, 0.5% (v/v) in PBS) was injected intravenously into NGC-N-infected STAT1^{-/-} mice at day 8 or 9 after infection; this was the point at which control mice began to exhibit signs of illness, such as ruffled fur and lethargy. At 2 h after injection, mice were killed and extensively perfused with PBS; tissues were harvested and weighed before the addition of formamide (2 ml of formamide per gram of tissue). Samples were incubated at 55 °C for 2.5 h, after which the Evans blue concentrations in formamide extracts were quantified by measuring A₆₁₀. Data are presented as A₆₁₀ per gram tissue weight. Blood samples were collected

from tail veins, and the levels of TNF- α and IP-10 and the titre of infectious virus in the sera were determined by ELISA (R&D Systems) and plaque assay in BHK-21 cells, respectively.

28. Chang, Y. C. *et al.* Modulation of macrophage differentiation and activation by decoy receptor 3. *J. Leukoc. Biol.* **75**, 486–494 (2004).
29. Hsu, T. L. *et al.* Modulation of dendritic cell differentiation and maturation by decoy receptor 3. *J. Immunol.* **168**, 4846–4853 (2002).
30. Oshiumi, H., Matsumoto, M., Funami, K., Akazawa, T. & Seya, T. TICAM-1, an adaptor molecule that participates in Toll-like receptor 3-mediated interferon- β induction. *Nature Immunol.* **4**, 161–167 (2003).

Single-stranded DNA-binding protein hSSB1 is critical for genomic stability

Derek J. Richard¹, Emma Bolderson¹, Liza Cubeddu^{2,3}, Ross I. M. Wadsworth², Kienan Savage^{1,4}, Girdhar G. Sharma⁵, Matthew L. Nicolette⁶, Sergie Tsvetanov¹, Michael J. McIlwraith⁷, Raj K. Pandita⁵, Shunichi Takeda⁸, Ronald T. Hay⁹, Jean Gautier¹⁰, Stephen C. West⁷, Tanya T. Paull⁶, Tej K. Pandita⁵, Malcolm F. White² & Kum Kum Khanna¹

Single-strand DNA (ssDNA)-binding proteins (SSBs) are ubiquitous and essential for a wide variety of DNA metabolic processes, including DNA replication, recombination, DNA damage detection and repair¹. SSBs have multiple roles in binding and sequestering ssDNA, detecting DNA damage, stimulating nucleases, helicases and strand-exchange proteins, activating transcription and mediating protein–protein interactions. In eukaryotes, the major SSB, replication protein A (RPA), is a heterotrimer¹. Here we describe a second human SSB (hSSB1), with a domain organization closer to the archaeal SSB than to RPA. Ataxia telangiectasia mutated (ATM) kinase phosphorylates hSSB1 in response to DNA double-strand breaks (DSBs). This phosphorylation event is required for DNA damage-induced stabilization of hSSB1. Upon induction of DNA damage, hSSB1 accumulates in the nucleus and forms distinct foci independent of cell-cycle phase. These foci colocalize with other known repair proteins. In contrast to RPA, hSSB1 does not localize to replication foci in S-phase cells and hSSB1 deficiency does not influence S-phase progression. Depletion of hSSB1 abrogates the cellular response to DSBs, including activation of ATM and phosphorylation of ATM targets

after ionizing radiation. Cells deficient in hSSB1 exhibit increased radiosensitivity, defective checkpoint activation and enhanced genomic instability coupled with a diminished capacity for DNA repair. These findings establish that hSSB1 influences diverse end-points in the cellular DNA damage response.

Ionizing radiation and anti-cancer drugs can induce DNA DSBs, which are highly cytotoxic lesions. In the S and G2 phases of the cell cycle, homologous recombination can be used to repair DSBs. To initiate homologous recombination, DNA is resected and then bound by RPA, a eukaryotic SSB, to facilitate Rad51 nucleofilament formation and strand invasion². Here we show that, in addition to RPA, the human genome encodes two further conserved SSB homologues, present on chromosomes 12q13.3 and 2q32.3, which we have named hSSB1 and hSSB2, respectively. The main focus of this study, hSSB1, is highly represented in EST libraries from a variety of tissues. It is conserved in metazoa, comprising an amino-terminal oligonucleotide/oligosaccharide-binding-fold domain, followed by a more divergent carboxy-terminal domain (Supplementary Fig. 1). Like RPA, recombinant hSSB1 binds specifically to ssDNA substrates (Fig. 1a and Supplementary Fig. 2), in particular to

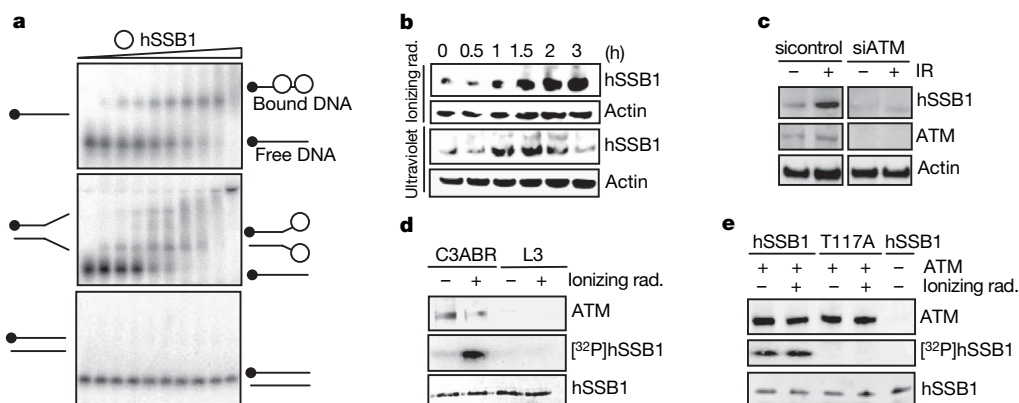


Figure 1 | ATM-dependent stabilization and phosphorylation of hSSB1 after ionizing radiation. **a**, Electrophoretic mobility shift analysis showing binding of recombinant hSSB1 to ssDNA substrates, d30T (top), a synthetic replication fork (middle) and dsDNA (bottom). The radiolabel is marked with a black circle. **b**, Immunoblots of hSSB1 using cell extracts from neonatal foreskin fibroblasts (NFFs) exposed to ionizing (6 Gy) or ultraviolet (20 J m⁻²) radiation. Cells were harvested at the indicated time

points and immunoblotted for hSSB1. **c**, Western blots of hSSB1 using ionizing radiation-treated (6 Gy) extracts from NFFs transfected with ATM siRNA. **d**, ATM was immunoprecipitated from mock or ionizing radiation (6 Gy)-treated normal (C3ABR) and A-T (L3) cell lines. *In vitro* kinase assays were performed using recombinant hSSB1 as a substrate. **e**, Phosphorylation of hSSB1 (number denotes the position of threonine residue substituted with alanine) by immunoprecipitated ATM kinase.

¹Signal Transduction Laboratory, Queensland Institute of Medical Research, Brisbane, Queensland 4029, Australia. ²Centre for Biomolecular Sciences, University of St Andrews, North Haugh, St Andrews, Fife KY16 9ST, UK. ³School of Molecular and Microbial Biosciences, University of Sydney, Sydney, New South Wales 2006, Australia. ⁴Central Clinical Division, School of Medicine, University of Queensland, Queensland 4072, Australia. ⁵Department of Radiation Oncology, Washington University School of Medicine, St Louis, Missouri 63108, USA. ⁶Department of Molecular Genetics and Microbiology, University of Texas at Austin, Austin, Texas 78712, USA. ⁷London Research Institute, Clare Hall Laboratories, Cancer Research UK, South Mimms, Hertfordshire EN6 3LD, U.K. ⁸Department of Radiation Genetics, Kyoto University Graduate School of Medicine, Kyoto 606-8501, Japan. ⁹Division of Gene Regulation and Expression, Wellcome Biocentre, University of Dundee, Dundee DD1 5EH, U.K. ¹⁰Institute for Cancer Genetics, Columbia University Medical Center, New York, New York 10032, USA.

polypyrimidines (Supplementary Fig. 3). The binding affinity increases significantly with the length of the DNA substrate (Supplementary Fig. 4).

RPA has several functions in the cell, including roles in DNA replication, recombination and repair¹. To investigate if hSSB1 functions in similar pathways, we analysed the response of hSSB1 to DNA damage. Cells were treated with ionizing or ultraviolet radiation, and hSSB1 was detected by western blotting (see Supplementary Fig. 5 for protein purification, short interfering RNA (siRNA)-mediated knockdown and antibody characterization). We found that hSSB1 accumulated in the cell in response to DNA damage (Fig. 1b), and that this was due to protein stabilization as treatment with the proteasome inhibitor MG132 led to a similar stabilization (Supplementary Fig. 6).

ATM kinase activity is essential for cellular signalling in response to DNA breaks³. siRNA-mediated depletion of ATM, inhibition of ATM activity (wortmannin treatment) or deficiency of ATM in ataxia telangiectasia (A-T) cells resulted in an inability to stabilize hSSB1 after ionizing radiation (Fig. 1c and Supplementary Fig. 7). Co-immunoprecipitation and glutathione S-transferase (GST)–ATM

fragment⁴ pull-down assays indicated that interaction between hSSB1 and ATM was direct and mediated by ATM fragment 4 (amino acids 772–1102) (Supplementary Figs 8 and 9).

ATM phosphorylates several proteins involved in the DNA damage-response pathway³. *In vitro* kinase assays⁵ showed that hSSB1 was a substrate of ATM (Fig. 1d) and that the phosphorylation site was mapped to T117 because mutation of the threonine to alanine blocked phosphorylation by ATM (Fig. 1e). To determine the functional relevance of this phosphorylation site *in vivo*, we compared the expression levels of exogenous GFP-hSSB1 wild-type and T117A mutant in HeLa cells. After ionizing radiation treatment, stabilization was only observed for the wild-type protein (Supplementary Fig. 10a), confirming the role of ATM in ionizing-radiation-induced stabilization. The stabilization defect is unlikely to be due to incorrect folding because recombinant T117A hSSB1 mutant folds correctly and shows similar ssDNA binding as wild-type hSSB1 (Supplementary Fig. 11a, b). Furthermore, GFP-hSSB1 and T117A were rapidly stabilized in the presence of MG132, whereas T117E mutant was not. This indicates that phosphorylation of hSSB1 at T117 prevents its degradation by the proteasome (Supplementary Fig. 10b).

After exposure to ionizing radiation, many repair proteins, including γ -H2AX, localize rapidly to sites of damage⁶. Immunofluorescence showed that hSSB1 localizes to prominent nuclear foci that formed within 30 min of DNA damage and persisted up to 8 h (Fig. 2a and data not shown). Foci for hSSB1 co-localized with most γ -H2AX foci after ionizing radiation and at a single *I-SceI*-induced DSB⁷ (Fig. 2a). Unlike hSSB1, we failed to observe significant co-localization of ionizing-radiation-induced RPA and γ -H2AX foci (Supplementary Fig. 12). Chromatin immunoprecipitation⁸ revealed that hSSB1 is present close (94–378 base pairs) to the *I-SceI*-induced DSB, suggesting that it might have a direct role in DSB repair (Fig. 2b).

The observation that hSSB1, like RPA, is recruited to sites of DSBs raises the issue of how these proteins are coordinated. RPA forms foci at sites of DNA replication in unperturbed S phase and after DNA damage both in S- and G2-phase cells^{9,10}. We examined whether hSSB1 exhibited similar foci formation kinetics and cell-cycle dependence. Before exposure to ionizing radiation, 22% of cells were positive for RPA foci, increasing to about 40% 3 h after treatment. By contrast, there were very few cells (less than 5%) positive for hSSB1 foci before exposure to ionizing radiation. However, 30 min after exposure, more than 95% of cells contained hSSB1 foci. Therefore, unlike RPA, hSSB1 foci formation is not cell-cycle dependent. In cells that contained both RPA and hSSB1 foci, we observed little co-localization (less than 5%), although it is interesting to note that about 27% of RPA foci were in close proximity to hSSB1 (less than 50 nm) (Fig. 2c). This proximity may represent different sub-compartments within one repair centre or a snapshot of the temporally dynamic composition of foci¹¹. The lack of direct co-localization indicates functional differences between these proteins in the repair process, although their proximity might suggest that they function at the same sites of repair.

The localization of hSSB1 to repair foci suggests it may act directly in DNA repair or in ionizing-radiation-induced signalling events. In response to DSBs, the cell cycle is arrested, before progression into S phase (G1/S checkpoint) or commitment to mitosis (G2/M checkpoint)³. These arrests are eventually released when the DNA lesions have been repaired. To determine whether hSSB1 was involved in checkpoint activation, we examined the effects of hSSB1 depletion on the activation of G1/S¹² and G2/M¹³ checkpoints after ionizing radiation. Unlike RPA¹⁴, depletion of hSSB1 by siRNA (sihSSB1) did not influence the number of cells in S phase or progression through it (Supplementary Fig. 13). Although control siRNA-treated cells arrest in G1 and G2 in response to ionizing radiation, hSSB1-depleted cells failed to arrest at either checkpoint (Fig. 3a and Supplementary Fig. 14). Cdc25a, a marker for rapid G1 arrest¹⁵, is normally degraded

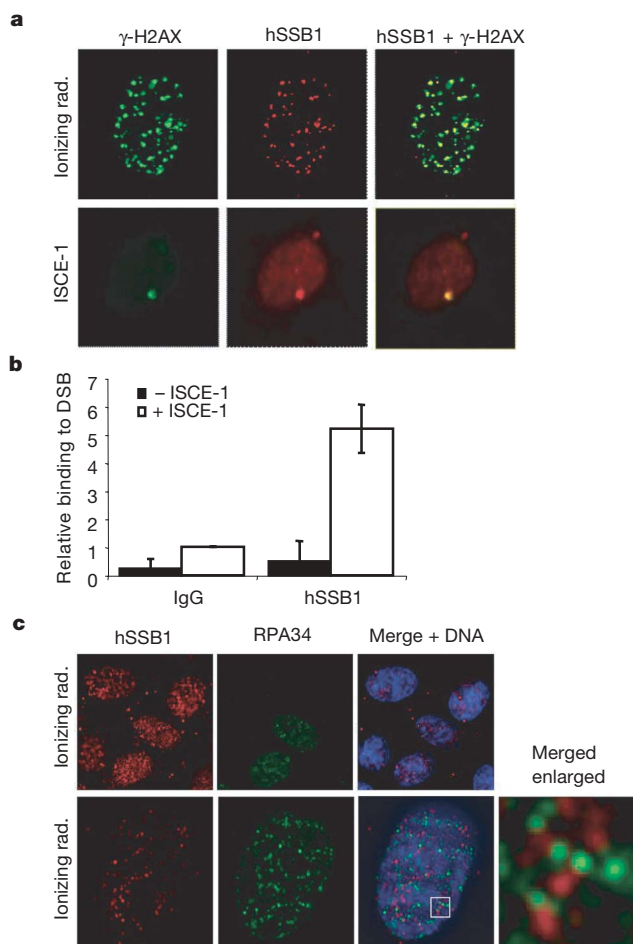


Figure 2 | hSSB1 localizes to DNA repair foci after ionizing radiation.

a, Irradiated (6 Gy) NFFs were extracted with detergent before fixation and stained with anti-hSSB1 and anti- γ H2AX antibodies (top panel; Leica TCS BI-15 microscope). hSSB1 and γ -H2AX co-localize at a single DSB induced by the *I-SceI* restriction enzyme in MCF7 DRGFP cells⁷ (bottom panel; Olympus BX61 microscope). **b**, ChIP analysis of hSSB1 on a unique DSB induced by *I-SceI* *in vivo*. Real-time PCR on ChIP samples used primers directed at 94–378 nucleotides from the DSB⁸. The enrichment of hSSB1 after induction of the DSB was compared with that of an IgG control (\pm s.d., $n = 3$). **c**, hSSB1 and RPA34 foci do not co-localize. NFFs were irradiated and 2 h later fixed and stained with anti-hSSB1 and anti-RPA34 antibodies. Images were acquired using Deltavision Personal DV. An enlargement of the merged image is displayed, demonstrating proximity of hSSB1 and RPA foci.

after ionizing radiation, but in hSSB1-deficient cells Cdc25a levels remained stable (Supplementary Fig. 15). The ATM-mediated phosphorylation of many checkpoint proteins is vital for ionizing-radiation-induced checkpoint activation³. To determine if this cascade was functional in the absence of hSSB1, we analysed the phosphorylation status of several ATM substrates in hSSB1-deficient cells. As expected, irradiation of control cells led to the autophosphorylation of ATM, phosphorylation of p53, Chk2, Chk1 and NBS1 (Fig. 3b). By contrast, hSSB1-deficient fibroblasts failed to show a similar degree of ATM autophosphorylation and phosphorylation of ATM targets after irradiation (Fig. 3b and Supplementary Fig. 16). Furthermore, defective phosphorylation of Chk1 also suggests a function of hSSB1 in ATR-dependent signalling.

Like hSSB1, the MRN complex plays a crucial role in the regulation of ATM activation and activity. We therefore determined whether hSSB1 affects MRN-dependent regulation of ATM activity¹⁶. We found that wild-type hSSB1 had no effect on ATM activity, whereas the phospho-mimic mutant consistently increased activity by more than twofold, 2.3 ± 0.71 (s.e.m., $n = 3$) (Fig. 3c). hSSB1 itself was not phosphorylated in this assay owing to the vast molar excess of p53. These results suggest that, once ATM is activated, a positive feedback loop is initiated through phosphorylated hSSB1, which in turn reinforces ATM activation. ATM signalling was examined in cells expressing GFP-hSSB1, T117A and T117E. Intriguingly, the expression of T117A suppressed ATM activation, as measured by S1981 phosphorylation, whereas the phospho-mimic T117E enhanced phosphorylation by approximately twofold (Fig. 3d). Together, these results indicate that hSSB1 amplifies ATM-dependent signalling.

Given that ATM is crucial for cell survival after ionizing radiation, we next examined the sensitivity of hSSB1-deficient cells to ionizing radiation. Cells deficient in hSSB1 displayed hypersensitivity to ionizing radiation (Fig. 4a). We were unable to assess cell survival after longer time points because of the severe death phenotype exhibited by ionizing radiation-treated hSSB1-deficient cells. Indeed, we usually observed greater than 90% cell death within 96 h of exposure to very low doses of radiation (0.5 Gy). Furthermore, we found that

even depletion of hSSB1 by 50% resulted in ionizing radiation sensitivity (Supplementary Fig. 17). Cells deficient in hSSB1 also displayed higher frequencies of chromosomal aberrations (chromosome and chromatid breaks, with fragments and telomere fusions) after exposure to ionizing radiation (Fig. 4b and Supplementary Fig. 18). These findings demonstrate that hSSB1 plays a functionally important role in allowing cells to repair genotoxic damage and maintain chromosome stability.

Because SSBs are required for homologous recombination¹⁷, we determined whether hSSB1 was important for homologous recombination in human cells. Depletion of hSSB1 from the MCF7 cell line, stably integrated with a GFP-based homologous recombination reporter construct^{7,18}, resulted in an approximate fivefold reduction in homologous recombination, as measured by GFP-positive cells after *I-SceI* expression (Supplementary Fig. 19). Rad51 binds ssDNA at resected DSBs to form nucleoprotein filaments that mediate the invasion of ssDNA into homologous duplex DNA². These reactions are stimulated by the presence of RPA^{19,20}. We therefore compared the ability of RPA and hSSB1 to stimulate Rad51-mediated strand invasion. We found that hSSB1 stimulated Rad51-mediated D-loop formation to a similar extent as RPA (Fig. 4c). Wild-type and T117E hSSB1 showed comparable stimulation in this assay, indicating that ATM-dependent hSSB1 phosphorylation does not appear to regulate strand invasion. Like RPA, hSSB1 was unable to stimulate strand invasion when added before Rad51 (data not shown). The stimulation of Rad51-mediated strand invasion by hSSB1 is likely to occur by a mechanism similar to that described previously with RPA^{21,22}. Consistent with these findings, Rad51 was not recruited to foci after ionizing radiation in hSSB1-deficient cells (Supplementary Fig. 20). Rad51 and hSSB1 were found to co-localize in 12% of cells treated with ionizing radiation, an additional 14% of foci were proximal (less than 50 nm) (Supplementary Fig. 21). An interaction between hSSB1 and Rad51 was also detected by co-immunoprecipitation (Supplementary Fig. 22).

In summary, we have identified a novel protein, hSSB1, that plays a key role in DNA damage response. Before this study, RPA was

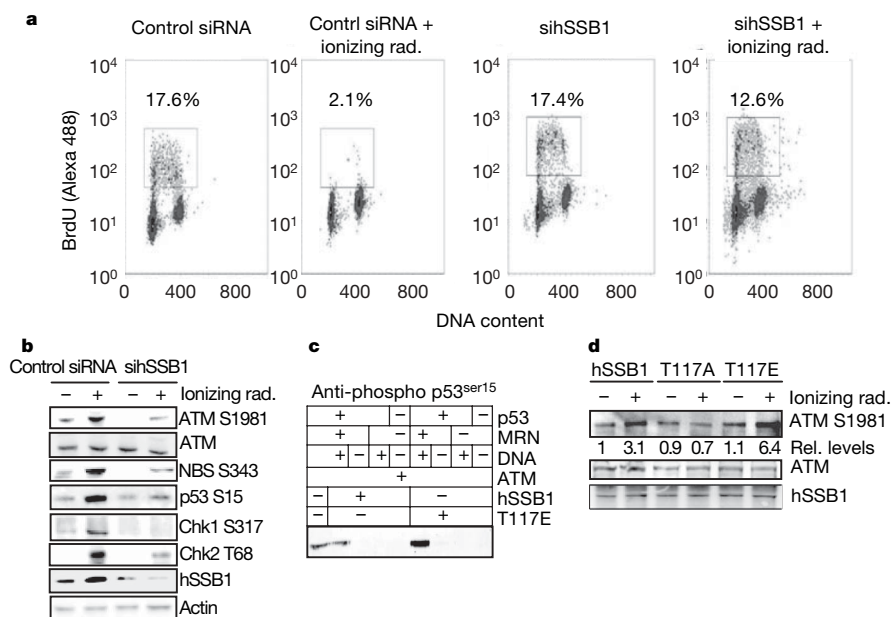
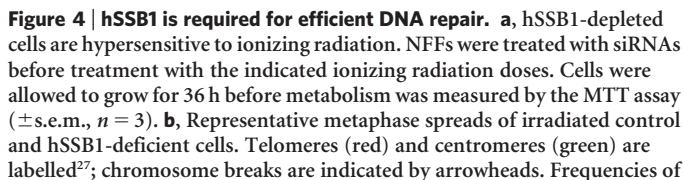


Figure 3 | hSSB1 inhibition results in defective ionizing-radiation-induced checkpoint activation and defective ATM signalling. **a**, hSSB1 inhibition results in defective ionizing-radiation-induced G1/S checkpoint. NFFs were mock-treated or irradiated (6 Gy) and labelled with BrdU, stained with anti-BrdU-Alexa488 antibodies and propidium iodide. The percentage of BrdU-positive cells (boxed area) was determined by fluorescence-activated cell sorting. **b**, hSSB1 is required for ionizing-radiation-induced ATM activation and activity. SiRNA-transfected NFFs were treated with ionizing radiation

and subjected to immunoblotting with the antibodies indicated. **c**, hSSB1-T117E affects MRN-dependent ATM activity. Purified ATM was incubated with DNA and MRN in the presence of wild-type hSSB1 or T117E mutant; phosphorylation of GST-p53 was detected using anti-Ser15 p53 antibody. **d**, hSSB1-T117E induces hyperactivation of ATM. HeLa cells transiently transfected with the indicated constructs were treated with ionizing radiation and immunoblotted with the antibodies indicated.



1. Wold, M. S. Replication protein A: a heterotrimeric single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu. Rev. Biochem.* **66**, 3050–3059 (1997).
2. West, S. C. Molecular views of recombination proteins and their control. *Nature Rev. Mol. Cell Biol.* **4**, 435–445 (2003).
3. Khanna, K. K. & Jackson, S. P. DNA double-strand breaks: signaling, repair and the cancer connection. *Nature Genet.* **27**, 247–254 (2001).
4. Khanna, K. K. *et al.* ATM associates with and phosphorylates p53: mapping the region of interaction. *Nature Genet.* **20**, 398–400 (1998).
5. Gatei, M. *et al.* Role for ATM in DNA damage-induced phosphorylation of BRCA1. *Cancer Res.* **60**, 3299–3304 (2000).
6. Bekker-Jensen, S. *et al.* Spatial organization of the mammalian genome surveillance machinery in response to DNA strand breaks. *J. Cell Biol.* **173**, 195–206 (2006).
7. Pierce, A. J., Johnson, R. D., Thompson, L. H. & Jasin, M. XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev.* **13**, 2633–2638 (1999).
8. Rodrigue, A. *et al.* Interplay between human DNA repair proteins at a unique double-strand break *in vivo*. *EMBO J.* **25**, 222–231 (2006).
9. Jazayeri, A. *et al.* ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. *Nature Cell Biol.* **8**, 37–45 (2006).
10. Soutoglou, E. *et al.* Positional stability of single double-strand breaks in mammalian cells. *Nature Cell Biol.* **9**, 675–682 (2007).
11. Lisby, M., Barlow, J. H., Burgess, R. C. & Rothstein, R. Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins. *Cell* **16**, 328–334 (2004).

12. Fabbro, M. *et al.* BRCA1-BARD1 complexes are required for p53Ser-15 phosphorylation and a G1/S arrest following ionising radiation-induced DNA damage. *J. Biol. Chem.* **279**, 31251–31258 (2004).
13. Xu, B., Kim, S. T. & Kastan, M. B. Involvement of Brca1 in S-phase and G(2)-phase checkpoints after ionising irradiation. *Mol. Cell. Biol.* **21**, 3445–3450 (2001).
14. Dodson, G. E., Shi, Y. & Tibbetts, R. S. DNA replication defects, spontaneous DNA damage, and ATM-dependent checkpoint activation in replication protein A-deficient cells. *J. Biol. Chem.* **279**, 34010–34014 (2004).
15. Mailand, N. *et al.* Rapid destruction of human Cdc25A in response to DNA damage. *Science* **288**, 1425–1429 (2000).
16. Lee, J. H. & Paull, T. T. ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. *Science* **308**, 551–554 (2005).
17. Wang, Y. *et al.* Mutation in *Rpa1* results in defective DNA double-strand break repair, chromosomal instability and cancer in mice. *Nature Genet.* **37**, 750–755 (2002).
18. Zhang, J., Ma, Z., Treszezamsky, A. & Powell, S. N. MDC1 interacts with Rad51 and facilitates homologous recombination. *Nature Struct. Mol. Biol.* **12**, 902–909 (2005).
19. McIlwraith, M. J. *et al.* Reconstitution of the strand invasion step of double-strand break repair using human RAD51, RAD52 and RPA proteins. *J. Mol. Biol.* **304**, 151–164 (2000).
20. Song, B. W. & Sung, P. Functional interactions among yeast Rad51 recombinase, Rad52 mediator, and replication protein A in DNA strand exchange. *J. Biol. Chem.* **275**, 15895–15904 (2000).
21. New, J. H. & Kowalczykowski, S. C. Rad52 protein has a second stimulatory role in DNA strand exchange that complements replication protein-A function. *J. Biol. Chem.* **277**, 26171–26176 (2002).
22. Sugiyama, T. & Kowalczykowski, S. C. Rad52 protein associates with replication protein A (RPA)-single-stranded DNA to accelerate Rad51-mediated displacement of RPA and presynaptic complex formation. *J. Biol. Chem.* **277**, 31663–31672 (2002).
23. Wadsworth, R. I. & White, M. F. Identification and properties of the crenarchaeal single-stranded DNA binding protein from *Sulfolobus solfataricus*. *Nucleic Acids Res.* **29**, 914–920 (2001).
24. Bartkova, J. *et al.* DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature* **434**, 864–870 (2005).
25. Aparicio, O. *et al.* in *Current Protocols in Molecular Biology* (eds Ausubel, F. A. *et al.*) 21.3.1–21.3.17 (John Wiley, New York, 2005).
26. Slavotinek, A., McMillan, T. J. & Steel, C. M. Measurement of radiation survival using the MTT assay. *Eur. J. Cancer* **30**, 1376–1382 (1994).
27. Pandita, R. K. *et al.* Mammalian Rad9 plays a role in telomere stability, S- and G2-phase-specific cell survival, and homologous recombinational repair. *Mol. Cell. Biol.* **26**, 1850–1864 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This study was supported by grants from the National Health and Medical Research Council of Australia (to K.K.K.), Cancer Research UK (to S.C.W.), the Biotechnology and Biological Sciences Research Council of the UK (to M.F.W.), the National Institutes of Health (CA92245 to J.G.; CA10445 and CA123232 to T.K.P.) and the American Cancer Society (RGS-04-173-01-CCG to T.T.P.).

Author Information The genes hSSB1 and hSSB2 are deposited in NCBI under accession numbers NM_024068 and NM_001031716, respectively. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to K.K.K. (kumkumK@qimr.edu.au) or M.F.W. (mfw2@st-and.ac.uk).

METHODS

Plasmids and siRNA. Full-length *hSSB1* was amplified by PCR from HeLa-cell complementary DNA and cloned in the *Bam*HI and *Sal*I sites of bacterial expression His-tag vector pET28c; *hSSB2* was cloned into bacterial expression His-tag vector pET19b. GFP-*hSSB1* fusion protein was cloned into the *Hind*III and *Kpn*I sites of pEGFP-C1, and threonine 117 was mutated to alanine using QuickChange site-directed mutagenesis, as described previously¹².

siRNAs were chemically synthesized (Invitrogen) with a two-nucleotide deoxythymidine overhang at the 3'-end. Individual siRNA sequences were as follows: *hSSB1* (sense) 5'-GACAAAGGACGGGCAUGAGdTdT, (antisense) 5'-CUCAUGCCCGUCCUUUGUCdTdT, *ATM* (sense) 5'-GCGCCUGAUUCG-AGAUCdUdTdT, (antisense) 5'-AGGAUCUCGAAUCAGGCGCdTdT. Cells were transfected twice with siRNA at a 24 h interval and processed 48 h later.

The 100-base oligonucleotide used in the D-loop assay was: 5'-GGGCGAA-TTGGGCCCGACGTCGCATGCTCCTCTAGACTCGAGGAATTCGGTACCCC-GGGTTCGAAATCGATAAGCTTACAGTCTCCATTAAAGGACAAG-3'. The oligo was 5'-³²P-end-labelled using polynucleotide kinase (New England Biolabs) and [γ -³²P]ATP.

Purification of recombinant protein. BL21 cells expressing His-tagged *hSSB1* were lysed in buffer A (20 mM Tris pH 8.0, 50 mM NaCl and 30 mM imidazole) with the addition of protease inhibitor cocktail (Complete Mini EDTA-free Protease, Roche) and 0.1% Triton X-100. Clarified soluble *E. coli* cell lysate pretreated with DNaseI (30 min, 20 °C) was applied to a Hi-Trap metal chelating column (GE Healthcare) loaded with nickel in buffer A and eluted using a gradient of 0–100% 300 mM imidazole. *hSSB1* protein isolated by nickel chromatography was diluted 3× in buffer A with 1 mM DTT and no NaCl, and loaded onto a heparin column (HiTrap, GE Healthcare). Protein was eluted using a gradient of 0–100% 1 M NaCl. *hSSB1* was concentrated and re-applied to the gel filtration column (Superdex 200) in buffer B (20 mM Tris pH 8.0, 150 mM NaCl and 1 mM DTT). Fractions were finally pooled, concentrated, snap frozen and stored at –80 °C.

Electrophoretic mobility shift analysis. The interaction of *hSSB1* with d30T oligonucleotide was investigated using native acrylamide electrophoretic mobility shift analysis. Increasing concentrations of *hSSB1* were incubated with ³²P-labelled d30T ssDNA (50 pmol) in buffer (20 mM HEPES, pH 7.3, 100 mM KCl and 1 mM MgCl₂, 1 mg μ l^{–1} bovine serum albumin) at 20 °C for 30 min in 10 μ l total volume. Reactions were resolved on 10% native acrylamide/TBE gel. Gels were exposed to a phosphorimager plate and the image visualized with a Fuji FLA-5000 PhosphorImager.

Isothermal titration calorimetry. Calorimetric experiments used a VP-ITC instrument (MicroCal). All solutions were degassed before use. *hSSB1* samples were dialysed extensively against 20 mM HEPES buffer, pH 7.3, 100 mM KCl and 1 mM MgCl₂. Oligonucleotides were also dissolved in isothermal titration calorimetry (ITC) buffer. Binding experiments were performed in triplicate at 37 °C. A 370- μ l syringe, stirring at 300 r.p.m., was used to titrate the oligonucleotide into the sample cell containing approximately 1.4 ml *hSSB1*. Each titration consisted of a preliminary 1- μ l injection followed by up to 25 subsequent 10- μ l injections. Heats of dilution were measured in corresponding blank titrations by adding oligonucleotide to ITC buffer and/or ITC buffer to protein, and were found to be similar to heats observed at the end of protein–DNA titrations. ITC-binding isotherms were analysed using a simple single-binding site model with ITC data analysis software (ORIGIN) provided by the manufacturer.

Antibodies and immunofluorescence. Antibodies used in this study were supplied by Calbiochem (Rad50, Mre11, Rad51), Santa Cruz (Cdc25a), Upstate (γ H2AX), Roche (BRDU), Cell Signalling Technologies (pT68-Chk2, pS317-Chk1, pS15-p53), Merck (H3 S10) and Invitrogen (GFP and Alexa secondary antibodies). Sheep antiserum to *hSSB1* was raised against full-length recombinant His-tagged *hSSB1*. For immunofluorescent staining, cells were permeabilized with 20 mM HEPES (pH 8), 20 mM NaCl, 5 mM MgCl₂, 1 mM ATP, 0.1 mM N₂O₉, 1 mM NaF and 0.5% NP40 for 15 min on ice before fixation in 4% paraformaldehyde (w/v) in PBS for 10 min.

D-loop assay. For D-loop assay, reactions contained 5'-³²P-end-labelled 100-base ssDNA (1 μ M), and the indicated concentration of Rad51, *hSSB1* or RPA in standard buffer (25 mM Tris-acetate pH 7.5, 5 mM CaCl₂, 2 mM ATP, 1 mM DTT, 100 μ g μ l^{–1} bovine serum albumin). Proteins were added and mixed in the sequence indicated. After 5 min at 37 °C, excess supercoiled pPB4.3 DNA (0.3 mM) was added and incubated for 10 min. The products were deproteinized by the addition of one-fifth volume of stop buffer (0.1 M Tris-HCl pH 7.5, 0.1 M MgCl₂, 3% SDS and 10 mg μ l^{–1} proteinase K) followed by electrophoresis, autoradiography and quantification using a Typhoon Trio PhosphorImager (GE Healthcare).

LETTERS

HP1- β mobilization promotes chromatin changes that initiate the DNA damage response

Nabieh Ayoub¹, Anand D. Jeyasekharan¹, Juan A. Bernal¹ & Ashok R. Venkitaraman¹

Minutes after DNA damage, the variant histone H2AX is phosphorylated by protein kinases of the phosphoinositide kinase family, including ATM, ATR or DNA-PK¹. Phosphorylated (γ)-H2AX—which recruits molecules that sense or signal the presence of DNA breaks, activating the response that leads to repair^{2,3}—is the earliest known marker of chromosomal DNA breakage. Here we identify a dynamic change in chromatin that promotes H2AX phosphorylation in mammalian cells. DNA breaks swiftly mobilize heterochromatin protein 1 (HP1)- β (also called CBX1), a chromatin factor bound to histone H3 methylated on lysine 9 (H3K9me). Local changes in histone-tail modifications are not apparent. Instead, phosphorylation of HP1- β on amino acid Thr 51 accompanies mobilization, releasing HP1- β from chromatin by disrupting hydrogen bonds that fold its chromodomain around H3K9me. Inhibition of casein kinase 2 (CK2), an enzyme implicated in DNA damage sensing and repair^{4–6}, suppresses Thr 51 phosphorylation and HP1- β mobilization in living cells. CK2 inhibition, or a constitutively chromatin-bound HP1- β mutant, diminishes H2AX phosphorylation. Our findings reveal an unrecognized signalling cascade that helps to initiate the DNA damage response, altering chromatin by modifying a histone-code mediator protein, HP1, but not the code itself.

We used an enhanced green fluorescent protein (EGFP)-HP1- β fusion protein^{7–9} to monitor changes in the dynamic behaviour of HP1- β after DNA breakage. In undamaged cells, fluorescence recovery after photobleaching (FRAP), using a range of laser-light intensities (Supplementary Fig. 1a–d), showed that EGFP-HP1- β exhibits distinct molecular mobilities in euchromatin compared with heterochromatin^{7,8}. Exposure of cells to ionizing radiation alters the dynamics of EGFP-HP1- β FRAP in a statistically significant manner (Fig. 1a). Recovery in both euchromatin and heterochromatin occurs more swiftly, as demonstrated by a change in the association constant, K , calculated from a single component exponential fit for the FRAP curves (Supplementary Fig. 1e). Similar changes are induced by the clastogenic drug etoposide (data not shown). Therefore, HP1- β mobilization is not solely a response to ionizing radiation.

EGFP-HP1- β mobilization after DNA damage in euchromatin, as well as heterochromatin, is evident using fluorescence loss in photobleaching (FLIP), which measures the exchange of fluorescent molecules between distinct regions within a cell (Supplementary Fig. 1f), and thus, their relative diffusional mobility, provided that imaging and bleaching parameters are held constant¹⁰. The changes in FLIP dynamics of EGFP-HP1- β after exposure to ionizing radiation are equivalent in euchromatin and heterochromatin, with the best-fit

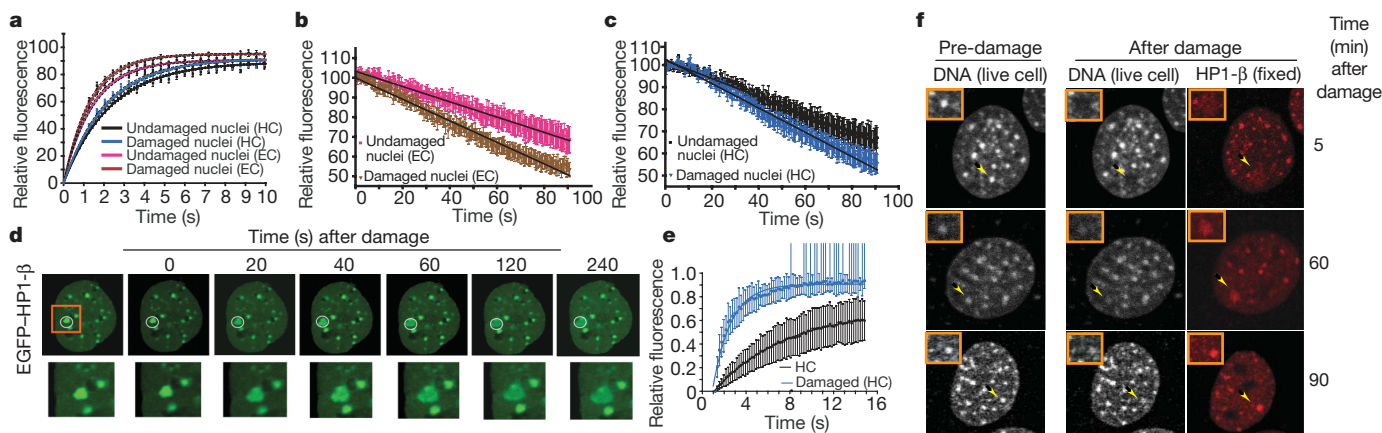


Figure 1 | Alterations in HP1- β dynamics and localization accompany its mobilization after DNA damage. **a**, Representation of the FRAP recovery curves ($n = 7$) of EGFP-HP1- β in euchromatin/nucleoplasm (EC) or heterochromatin (HC), before and after DNA damage by 10 Gy ionizing radiation, fitted (solid lines) to a single component exponential (dotted lines show 95% confidence intervals). Relative fluorescence intensity is plotted (y axis) against time (x axis). **b**, **c**, FLIP decay curves ($n = 7$) for euchromatin and heterochromatin before and after DNA damage, defining the acceleration in mobility of the EGFP-HP1- β protein after damage. The solid lines represent a single order polynomial fit of the data. **d**, Changes in the abundance and distribution of EGFP-HP1- β immediately after the

induction of laser-induced DNA damage in a single heterochromatin focus (white circle). The area bounded by the orange square is enlarged below each panel. **e**, FRAP curves for EGFP-HP1- β in heterochromatin before and after DNA damage ($n = 10$). **f**, Changes in endogenous HP1- β at defined time points (rows) after the induction of DNA damage to a heterochromatin focus. Hoechst staining of a live cell immediately before, and after, laser damage is shown in the first two columns. The third column shows the same cell after fixation and immunostaining for HP1- β . The area marked by the yellow arrow has been enlarged at the upper left corner of each frame. All results are typical of >3 independent experiments. Where shown, error bars represent s.e.m.

¹The Medical Research Council Cancer Cell Unit, Hutchison/MRC Research Centre, Hills Road, Cambridge CB2 0XZ, UK.

decay constant increasing from 0.4 ± 0.01 (s.e.m.) to 0.55 ± 0.01 in each ($P < 0.05$, $n = 7$, t -test), confirming exchange between the two compartments and suggesting that a similar mechanism causes mobilization in both (Fig. 1b, c).

We precisely targeted DNA double-strand breaks (DSBs), marked by γ -H2AX formation and 53BP1 recruitment, to specific living cells

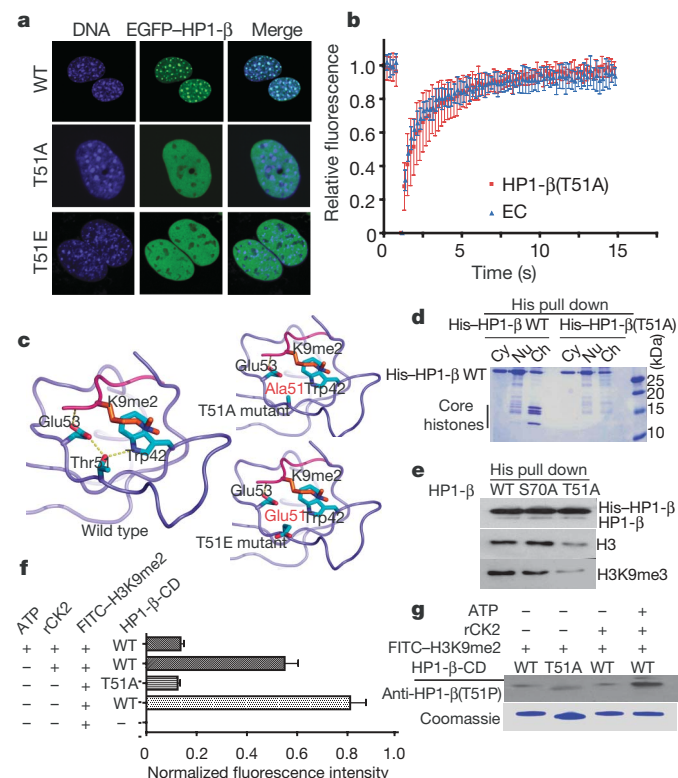


Figure 2 | Thr 51 substitution or phosphorylation alter the localization and dynamics of HP1-β by disrupting a hydrogen-bond network essential for the HP1-H3K9me interaction. **a**, Localization of wild-type (WT) EGFP-HP1-β (top panel) or mutant forms in which Thr 51 in the chromodomain has been replaced with alanine (T51A, middle panel) or glutamic acid (T51E, bottom panel). **b**, The dynamic behaviour of EGFP-HP1-β(T51A) measured by FRAP is similar to the high mobility of EGFP-HP1-β in euchromatin regions. Fluorescence recovery (relative fluorescence intensity, y axis) after photobleaching is shown over time ($n = 10$). Error bars represent s.d. **c**, A model for the interaction of the wild-type human HP1-β chromodomain bound to methylated K9 of histone H3 (K9me2), based on the PDB coordinates 1KNA, compared with its T51A and T51E mutants. A network of hydrogen bonds between the side chains of Glu 53, Thr 51 and Trp 42 of the human HP1-β chromodomain, marked as yellow dashed lines, is disrupted in the T51A and T51E mutants. **d**, Cytoplasmic (Cy), nuclear soluble (Nu) or chromatin-bound (Ch) proteins pulled down with His-tagged HP1-β wild-type or T51A mutant proteins in a Coomassie-stained 12% SDS-PAGE gel. The identity of the four core histones, pulled down by the wild-type but not the T51A mutant protein, was confirmed by mass spectrometry (data not shown). **e**, Western blot analysis of proteins pulled down with His-tagged HP1-β wild-type, S70A or T51A mutant proteins in nuclear extracts. Blots were probed with antibodies against HP1-β (top row), histone H3 (middle row) or histone H3K9me3 (bottom row). The S70A mutant, whose localization and dispersal after DNA damage are indistinguishable from the wild-type protein, serves as a control. All three variants equally pull down endogenous HP1-β. **f**, **g**, Interaction of the HP1-β chromodomain with methylated H3K9 is disrupted by CK2 phosphorylation on Thr 51. **f**, Fluorescence intensities (x axis) of a fixed volume of glutathione beads used to pull down a FITC-H3K9me2 peptide incubated with the different GST-HP1-β chromodomain fusion proteins listed ($n = 3$, error bars represent s.e.m.). Decreased peptide binding (that is, fluorescence intensity) occurs when the GST-HP1-β chromodomain is phosphorylated on Thr 51, as demonstrated in **g** by probing protein eluted from the glutathione beads with anti-HP1-β(T51P). Equal loading of GST-HP1-β chromodomain in each track is shown by Coomassie staining.

or discrete regions within a single nucleus (Supplementary Fig. 2a–c), using a 405-nm laser¹¹. Damage to a single heterochromatin focus (Fig. 1d, circled white) swiftly disperses EGFP-HP1-β, inciting local rearrangement and spreading within 20 s (Fig. 1d and Supplementary Movies 1 and 2), but with little overall variation in DNA staining (Supplementary Fig. 2d), distinct from reported changes in GFP-tagged histone H2B¹². The FRAP dynamics of EGFP-HP1-β changes locally at damage sites (Fig. 1e). Endogenous HP1-β visualized by immunofluorescence is also rearranged and dispersed from a damaged heterochromatin focus within 5 min after DNA damage, to re-accumulate gradually over 90 min (Fig. 1f), consistent with a physiological response to DNA breakage and repair.

HP1-β localizes to chromatin via its direct interaction with H3K9me (Supplementary Fig. 3a). Unexpectedly, neither the distribution nor abundance of H3K9me3 staining is altered after DNA damage. This is also true for trimethylation of histone H3 on lysine 4 (H3K4me3) (Supplementary Fig. 3b), and acetylation of histone H3 on lysines 14 or 18 (ref. 13) or acetylation of histone H4 on lysine 16 (Supplementary Fig. 3c). Thus, it seems unlikely that alterations in the histone code mobilize HP1-β after DNA damage. Instead, mobilization may result from changes in HP1-β itself, akin to previous reports^{14,15}.

Substitution of eight putative phosphorylation sites in HP1-β affected neither the localization nor the damage-induced dispersal of EGFP-HP1-β from heterochromatic foci (Supplementary Fig. 4). By contrast, alteration of Thr 51 in the HP1-β chromodomain either to Ala or to Glu (respectively expected to preclude, or mimic, phosphorylation) suffices to distribute mutant EGFP-HP1-β diffusely throughout the nucleus (Fig. 2a) and to confer rapid dynamics in FRAP (Fig. 2b). Similarly, replacement of the corresponding residue, Thr 50, in human HP1-α (Supplementary Fig. 5a) excludes HP1-α from heterochromatin (Supplementary Fig. 5b).

Thr 51 is evolutionarily conserved (Supplementary Fig. 6a). The structure of the HP1-β chromodomain (Fig. 2c), modelled from a *Drosophila melanogaster* HP1-H3K9me complex¹⁶, reveals that the hydroxyl moiety in the Thr 51 side chain participates in a hydrogen-bond network essential for complex formation. Phosphorylation or replacement of Thr 51 is expected to disrupt this hydrogen-bond network (Fig. 2c), diminishing the capacity of HP1-β to bind to chromatin via H3K9me3. Indeed, whereas His-tagged HP1-β (His-HP1-β) effectively pulls down the four core histones from cell extracts (Fig. 2d), mutant His-HP1-β(T51A) does not. Moreover, His-HP1-β and its Ser70Ala mutant both bind efficiently to histone H3K9me3 from chromatin, although mutant His-HP1-β(T51A) does not (Fig. 2e). Thus, Thr 51 modification suffices to release HP1-β from chromatin, although additional, DNA-damage-induced changes could also assist.

In human HP1-β, Thr 51 lies within an atypical phospho-acceptor site for CK2 (Supplementary Fig. 6b)¹⁷, divergent from the consensus Ser/Thr-X-X-Asp/Glu. Indeed, recombinant (r) CK2 can phosphorylate the wild-type HP1-β chromodomain (but not its T51A mutant form), confirming that this residue is a specific target (Supplementary Fig. 6c–e). An antibody specific to the phosphopeptide SDEDN(pThr)WEPEEC reacts *in vitro* with recombinant HP1-β wild-type protein solely after exposure to rCK2, but not with the T51A mutant (Supplementary Fig. 6f). *In vivo*, the antibody reacts with the wild-type but not the mutant protein (Supplementary Fig. 7a). This reactivity increases after etoposide-induced DNA damage, with kinetics parallel to γ -H2AX formation.

Phosphorylation of the HP1 chromodomain on Thr 51 by rCK2 releases it from binding to a fluorescein-tagged H3K9me2 peptide (Fig. 2f). The fluorescent peptide is pulled down by the HP1 chromodomain, but not by its T51A mutant. Thr 51 phosphorylation by rCK2 (Fig. 2g) diminishes binding of the fluorescent H3K9me2 peptide to background levels.

In cells treated with isoform-specific short interfering (si)RNAs targeting HP1-β, diminished HP1-β expression corresponds with

diminished anti-HP1- β (T51P) staining after etoposide-induced DNA damage (Fig. 3a, b). Ionizing radiation also elicits an increase in anti-HP1- β (T51P) staining, which is inhibited by pre-incubation with the phosphopeptide immunogen (Supplementary Fig. 7b). There is increased nucleoplasmic and local staining for HP1- β (T51P) specifically in damaged nuclei after γ -H2AX phosphorylation, and not in undamaged ones, after laser irradiation at 405 nm (Fig. 3c) or 365 nm (Supplementary Fig. 7c). Laser damage to a heterochromatin focus triggers dispersal of endogenous HP1- β accompanied by spreading HP1- β (T51P) staining around the damaged region (Fig. 3d). Together, these findings suggest that HP1- β (T51P) is created locally in damaged chromatin and thereafter released, mirroring the pattern of HP1- β rearrangement and dispersal.

Inhibition of CK2 activity by chemical or genetic means suppresses damage-induced changes in HP1- β dynamics and its phosphorylation on Thr 51. In cell extracts, anti-HP1- β (T51P) detects a species

with mobility corresponding to HP1 (Supplementary Fig. 7d), enhanced in a dose-dependent manner by etoposide treatment. Chemical inhibition of CK2 with 4,5,6,7-tetrabromo-benzimidazole (TBB) at 75 μ M for 6 h, a concentration required to suppress CK2 activity in living cells¹⁸, diminishes both the background detection and etoposide-induced enhancement of HP1- β (T51P) levels (Supplementary Fig. 7b, d). Genetic suppression of CK2 activity by an overexpressed dominant-negative form of the enzyme¹⁹ is enough to suppress Thr 51 phosphorylation after DNA damage (Fig. 4a, b). Finally, in living cells TBB suppresses the changes in FRAP dynamics (Fig. 4c) usually triggered by DNA damage. Together, these observations demonstrate that CK2 mediates the mobilization of HP1- β from chromatin after DNA breakage via Thr 51 phosphorylation.

Chromatin alterations may accompany phosphoinositide kinase activation^{20,21} or modulate γ -H2AX deposition along chromosomes^{12,22,23}, but whether they help to trigger the response to DNA breaks is unknown. We not only find that CK2 inhibition suffices to diminish histone H2AX phosphorylation after DNA breakage

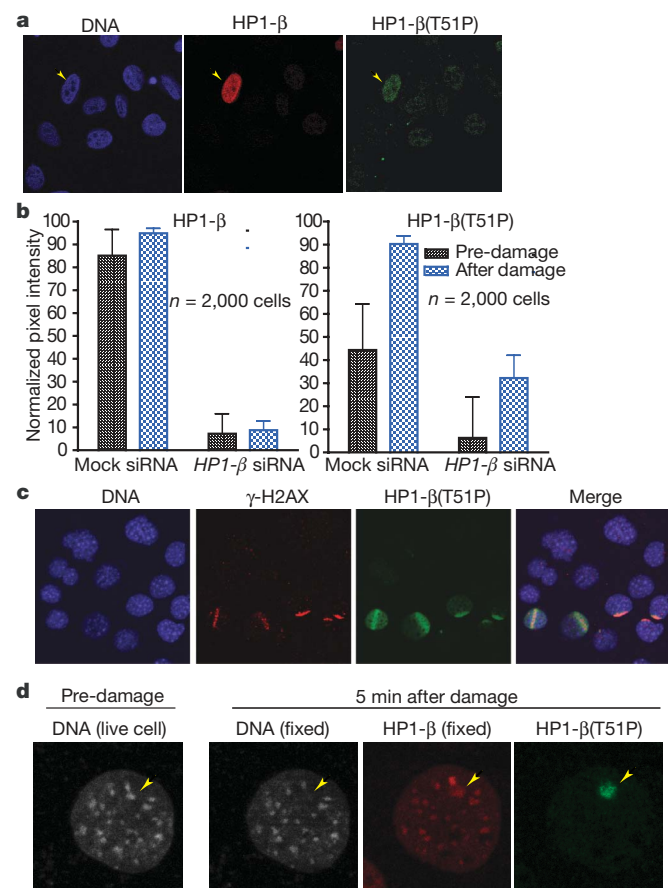


Figure 3 | Thr 51 phosphorylation is induced after DNA damage and accompanies HP1- β dispersal from damage sites. **a**, U2OS cells treated 72 h after HP1- β depletion with 20 μ M etoposide for 1 h before fixation. Nuclei (DNA in blue) were stained with anti-HP1- β (red) and anti-HP1- β (T51P) (green). HP1- β (T51P) staining is visible only in cells spared from HP1- β depletion (yellow arrow). **b**, Immunofluorescence intensity per cell for HP1- β (left) and HP1- β (T51P) (right) after HP1- β knockdown and etoposide treatment ($n = 2,000$, error bars represent s.d.) measured with a Cellomics HCS microscope. Both the induction of anti-HP1- β (T51P) staining after DNA damage in the mock siRNA control as well as a significant decrease of average HP1- β (T51P) staining intensity per cell after HP1- β depletion are evident. **c**, Representative field of mouse embryonic fibroblasts (MEFs) damaged by 405-nm laser micro-irradiation fixed and stained within 5 min after damage. HP1- β (T51P) staining (green) is present only in damaged nuclei (DNA in blue) marked by γ -H2AX formation (red). **d**, Localized induction of anti-HP1- β (T51P) (green) accompanying HP1- β dispersal (red) in a single 405-nm laser-damaged heterochromatin focus (yellow arrow). The first panel is a live cell image of Hoechst staining before radiation showing the undamaged heterochromatin focus.

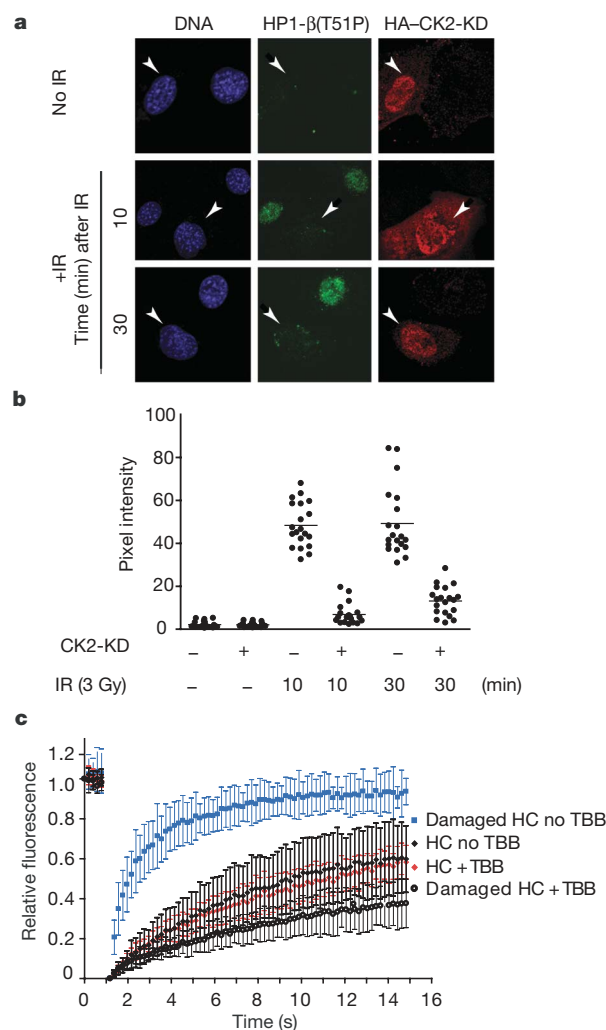


Figure 4 | Inhibition of CK2 suppresses the phosphorylation of HP1- β on Thr 51 and mobilization after DNA damage. **a**, MEFs (nuclei in blue) transfected with a haemagglutinin (HA)-epitope-tagged, kinase-dead (KD) K69M CK2 α' mutant. Anti-HA staining in red marks the transfected cells, whereas anti-HP1- β (T51P) staining is in green. Transfected cells (arrowheads) have less HP1- β (T51P) staining after 3 Gy ionizing radiation (IR) compared to neighbouring, untransfected cells. **b**, Quantitative representation of anti-HP1- β (T51P) staining intensity per cell under the specified experimental conditions ($n = 30$, $P < 0.05$ by analysis of variance (ANOVA)/Dunnett's post test). **c**, The dynamic behaviour of EGFP-HP1- β after DNA damage, with (black circles) or without (blue squares) pre-exposure to 75 μ M of the CK2 inhibitor TBB.

(Fig. 5a, b and Supplementary Fig. 8a, b), but also that immobilization of HP1- β on chromatin suppresses γ -H2AX formation. We fused EGFP-HP1- β to the histone H2B in a strategy previously used to render proteins constitutively chromatin-bound¹⁰ whether before or after DNA damage (Supplementary Fig. 9). The intensity of γ -H2AX staining after DNA damage is inhibited in cells expressing the immobilized H2B-EGFP-HP1- β fusion protein (Fig. 5c, d). Neither EGFP-HP1- β nor H2B-EGFP (Fig. 5c, d) expression to similar levels affects γ -H2AX staining, ruling out that this effect is merely the result of HP1- β overexpression or high levels of chromatin-bound H2B. Instead, our findings collectively show that mobilization of HP1- β from chromatin facilitates H2AX phosphorylation after DNA damage. We propose on this basis that dynamic alterations in chromatin structure triggered by HP1- β mobilization promote H2AX phosphorylation, so far the earliest known event in the response to chromosomal breakage.

How CK2 can be activated locally within seconds at damage sites is not understood. Intriguingly, although neither HP1- β nor CK2 seems to be a direct substrate of phosphoinositide-kinase-family kinases²⁴, chemical inhibition of these enzymes decreases but does not prevent HP1- β phosphorylation (Supplementary Fig. 10),

suggestive of crosstalk with the damage-activated CK2 signalling cascade.

Notably, HP1- β mobilization after DNA damage provides an unusual example of a mechanism that alters chromatin organization by targeting a histone-binding protein, rather than the histone code itself²⁵. Such a mechanism could facilitate both the rapid remodelling of chromatin after DNA damage, as well as its efficient reversal by HP1- β dephosphorylation.

METHODS SUMMARY

Transfection. We used established protocols for Lipofectamine 2000 (Invitrogen).

Microscopy and laser-induced DNA damage. DNA damage by laser-irradiation was by pre-sensitization with Hoechst 33342, followed by exposure to a 405-nm diode laser, or a pulsed 365-nm micropoint laser beam. DNA breakage was confirmed using staining for γ -H2AX and 53BP1. Dispersal of EGFP-tagged proteins was followed by live-cell sequential imaging every 0.3 s. Alternatively, dispersed proteins were detected by fixation and immunostaining. Dynamic exchange was studied by photobleaching EGFP-tagged proteins, and measuring fluorescence recovery in the bleached area (FRAP), or its loss from a distant area during repetitive bleach cycles (FLIP). All imaging of EGFP-tagged proteins was with an argon 488 laser line (attenuated to 1%) on a Zeiss LSM 510-META Confocor2. Average per-cell fluorescence intensity measurements were analysed with the Zeiss LSM software, from images acquired using identical acquisition parameters. Where indicated, immunofluorescence experiments were performed in 96-well plates and analysed using a Cellomics HCS automated microscope and associated software.

Biochemical assays. His-, glutathione S-transferase (GST)-, or antibody-mediated pull downs were performed using modifications of standard methods. Binding of a fluorescent methylated Lys 9 (K9me) peptide to GST-HP1 was quantified on a Fusion plate-reader.

Western blotting. Modifications to standard methods are described in the Methods section. Signal intensities were quantified by integrated density measurements on ImageJ software.

Statistical analysis. This was performed using GraphPad Prism 4.0.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 14 November 2007; accepted 28 February 2008.

Published online 27 April 2008.

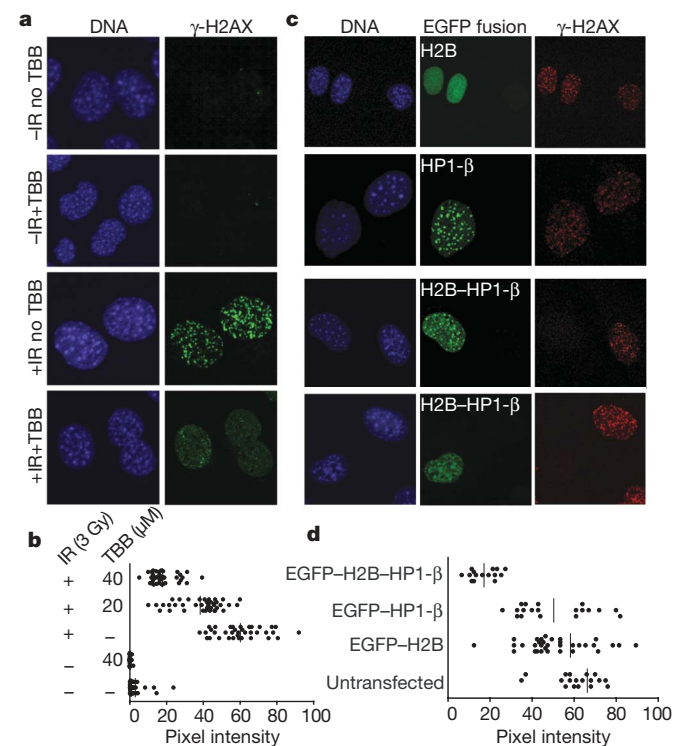


Figure 5 | HP1- β mobilization by CK2 promotes H2AX phosphorylation.

a, Pre-exposure of MEFs to 75 μ M TBB suppresses γ -H2AX foci formation at 5 min after 3 Gy ionizing radiation. DNA is stained blue; γ -H2AX is green. The top two rows show undamaged cells without (–IR no TBB) or with (–IR +TBB) pre-exposure; the bottom two rows represent ionizing-radiation-damaged cells without (+IR no TBB) or with TBB pre-exposure (+IR +TBB). **b**, Dot-plot showing that the pixel intensity of γ -H2AX staining after damage under identical imaging conditions is reduced in a dose-dependent manner by TBB pre-exposure ($n = 20$ per sample, $P < 0.01$, independent t -tests). **c**, An H2B-EGFP-HP1- β fusion protein constitutively immobilized on chromatin suppresses γ -H2AX foci formation at 5 min after 3 Gy ionizing radiation. DNA is stained blue, the different EGFP fusion proteins are stained green, and γ -H2AX is red. The top two rows show MEFs transfected with constructs encoding EGFP-HP1- β or H2B-EGFP, whereas the bottom two rows both show H2B-EGFP-HP1- β . Untransfected cells within the same field provide an internal control. **d**, Dot-plot representing the pixel intensities of γ -H2AX signal per cell under identical imaging conditions for cells with and without the transfected constructs ($n = 20$, $P < 0.05$ with t -test/ANOVA for multi-group comparison).

- Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. & Bonner, W. M. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.* **273**, 5858–5868 (1998).
- Fernandez-Capetillo, O., Lee, A., Nussenzweig, M. & Nussenzweig, A. H2AX: the histone guardian of the genome. *DNA Repair (Amst.)* **3**, 959–967 (2004).
- Lowndes, N. F. & Toh, G. W. DNA repair: the importance of phosphorylating histone H2AX. *Curr. Biol.* **15**, R99–R102 (2005).
- Allende-Vega, N., Dias, S., Milne, D. & Meek, D. Phosphorylation of the acidic domain of Mdm2 by protein kinase CK2. *Mol. Cell. Biochem.* **274**, 85–90 (2005).
- Cheung, W. L. *et al.* Phosphorylation of histone H4 serine 1 during DNA damage requires casein kinase II in *S. cerevisiae*. *Curr. Biol.* **15**, 656–660 (2005).
- Loizou, J. I. *et al.* The protein kinase CK2 facilitates repair of chromosomal DNA single-strand breaks. *Cell* **117**, 17–28 (2004).
- Festenstein, R. *et al.* Modulation of heterochromatin protein 1 dynamics in primary mammalian cells. *Science* **299**, 719–721 (2003).
- Cheutin, T. *et al.* Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science* **299**, 721–725 (2003).
- Schmiedeberg, L., Weisshart, K., Diekmann, S., Meyer Zu Hoerster, G. & Hemmerich, P. High- and low-mobility populations of HP1 in heterochromatin of mammalian cells. *Mol. Biol. Cell* **15**, 2819–2833 (2004).
- Lukas, C., Falck, J., Bartkova, J., Bartek, J. & Lukas, J. Distinct spatiotemporal dynamics of mammalian checkpoint regulators induced by DNA damage. *Nature Cell Biol.* **5**, 255–260 (2003).
- Rogakou, E. P., Boon, C., Redon, C. & Bonner, W. M. Megabase chromatin domains involved in DNA double-strand breaks *in vivo*. *J. Cell Biol.* **146**, 905–916 (1999).
- Kruhlak, M. J. *et al.* Changes in chromatin structure and mobility in living cells at sites of DNA double-strand breaks. *J. Cell Biol.* **172**, 823–834 (2006).
- Kuo, M. H. *et al.* Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature* **383**, 269–272 (1996).
- Zhao, T. & Eisenberg, J. C. Phosphorylation of heterochromatin protein 1 by casein kinase II is required for efficient heterochromatin binding in *Drosophila*. *J. Biol. Chem.* **274**, 15095–15100 (1999).
- Zhao, T., Heyduk, T. & Eisenberg, J. C. Phosphorylation site mutations in heterochromatin protein 1 (HP1) reduce or eliminate silencing activity. *J. Biol. Chem.* **276**, 9512–9518 (2001).

16. Jacobs, S. A. & Khorasanizadeh, S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* **295**, 2080–2083 (2002).
17. Formby, B. & Stern, R. Phosphorylation stabilizes alternatively spliced CD44 mRNA transcripts in breast cancer cells: inhibition by antisense complementary to casein kinase II mRNA. *Mol. Cell. Biochem.* **187**, 23–31 (1998).
18. Ruzzene, M., Penzo, D. & Pinna, L. A. Protein kinase CK2 inhibitor 4,5,6,7-tetrabromobenzotriazole (TBB) induces apoptosis and caspase-dependent degradation of haematopoietic lineage cell-specific protein 1 (HS1) in Jurkat cells. *Biochem. J.* **364**, 41–47 (2002).
19. Canton, D. A., Zhang, C. & Litchfield, D. W. Assembly of protein kinase CK2: investigation of complex formation between catalytic and regulatory subunits using a zinc-finger-deficient mutant of CK2 β . *Biochem. J.* **358**, 87–94 (2001).
20. Bakkenist, C. J. & Kastan, M. B. DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature* **421**, 499–506 (2003).
21. Ziv, Y. *et al.* Chromatin relaxation in response to DNA double-strand breaks is modulated by a novel ATM- and KAP-1 dependent pathway. *Nature Cell Biol.* **8**, 870–876 (2006).
22. Kim, J.-A., Kruhlak, M., Dotiwala, F., Nussenzweig, A. & Haber, J. E. Heterochromatin is refractory to γ H2AX modification in yeast and mammals. *J. Cell Biol.* **178**, 209–218 (2007).
23. Cowell, I. G. *et al.* γ H2AX foci form preferentially in euchromatin after ionising radiation. *PLoS ONE* **2**, e1057 (2007).
24. Matsuoka, S. *et al.* ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316**, 1160–1166 (2007).
25. Lomberk, G., Bensi, D., Fernandez-Zapico, M. E. & Urrutia, R. Evidence for the existence of an HP1-mediated subcode within the histone code. *Nature Cell Biol.* **8**, 407–415 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank L. Pellegrini for modelling the HP1–H3K9me2 structure; M. Daniels, S. Y. Peak-Chew, M. Lee, R. Kulkarni and P. Rowling for technical assistance; and T. Misteli, T. Jenuwein and D. Litchfield for gifts of material. N.A. acknowledges a fellowship from the International Agency for Research on Cancer, Lyon, France, J.A.B. a long-term fellowship from EMBO, Heidelberg, Germany, and A.D.J. a scholarship from the Gates Cambridge Trust. The UK Medical Research Council supports work in A.R.V.'s laboratory.

Author Contributions N.A. performed the experiments reported here, except that A.D.J. determined EGFP–HP1- β mobilization by FRAP and FLIP, demonstrated the release of Thr 51-phosphorylated HP1 from H3K9me, and performed the quantification of immunofluorescence, whereas J.A.B. helped to identify HP1- β phosphorylation sites, and measured the rCK2 kinetics. A.R.V. planned the project, helped to interpret the data and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.R.V. (arv22@cam.ac.uk).

METHODS

FRAP and FLIP assays. MEFs were seeded in 2-well Lab-Tek chambers (Nunc) after transfection with the EGFP-HP1- β construct. Before imaging, growth medium was replaced with phenol-red free L15 imaging medium. Live cell imaging and photobleaching experiments were performed using a $\times 40$ 1.2 NA water immersion lens on a Zeiss LSM 510 Meta confocal microscope (using a heating chamber to maintain the cells at 37 °C) with the 488-nm line of an argon laser operating at 6.1 A.

FRAP analysis was performed by a method previously established in our laboratory²⁶. For low-bleach FRAP (see Supplementary Fig. 1), selected regions of heterochromatin and euchromatin were simultaneously bleached in a diffraction-limited spot with a pulse of two iterations at 100% of the 15-mW 488-nm argon laser. Post-bleach images were acquired at a lower frequency of an image every 400 ms. Image bleach was corrected separately for euchromatin and heterochromatin regions using a nearby unbleached cell. For high-bleach FRAP experiments, photobleaching was carried out with 35 iterations at the same laser output, and images were collected every 290 ms. The normalized curves were plotted in GraphPad Prism, and recovery curves were fitted to a single order exponential equation $Y = Y_{\max}(1 - \exp(-KX))$ to yield a recovery constant K , shown to be different among groups by the sum of squares f -test at $P < 0.05$.

Transfected cells for analysis were chosen under epifluorescence. Transfection levels were controlled by using cells of similar nuclear size, that were visible but not saturated, at a fixed imaging intensity (1% transmission at maximal scan speed and fixed photomultiplier tube gain at an optical slice of 1.5 μm). An integrative time-bleach series was set up to take 5 initial images, then bleach a $2 \times 2 \mu\text{m}$ square with 30 iterations of 100% laser intensity in an euchromatin region every 4 images, for a total of 200 images. The cells were imaged with a fixed $\times 7$ zoom and fixed bleach region of interest (ROI) size in a 512×512 pixel image window to keep the time between images constant for different cells to allow comparison. Circular regions of heterochromatin and euchromatin of similar size (1- μm radius) equidistant at 10 μm from the site of bleaching were then chosen, and the pixel intensities during the FLIP procedure recorded. The procedure was performed on control and irradiated cells (10 min after ionizing radiation to allow temperature equilibration and decreased drift during scanning), with control cells also kept at room temperature for 20 min to compensate for the irradiation process. Nearby imaged—but unbleached—cells were used to correct for the effect of imaging during the FLIP procedure. For the data analysis, the raw pixel intensities were transferred into GraphPad Prism and converted into relative intensities as a percentage of the mean of the first three values. These intensities were plotted across time to generate the FLIP decay curves with standard errors being calculated from multiple measurements ($n = 7$, in repeated experiments, limited to avoid discrepancies from changes in time after irradiation). The curves were fit to a single order polynomial equation $y = a + bx$ and the decay constant b shown to be statistically different for different groups at $P < 0.05$ using the f -test.

Laser micro-irradiation. Laser micro-irradiation was used to induce DNA damage by modifying a previously described method²⁷. Briefly, cells were grown on chambered, 1-thickness borosilicate-glass cover slides (Nunc), incubated with $10 \mu\text{g ml}^{-1}$ Hoechst dye 33342 for 5 min at 37 °C, washed twice with Leibovitz's L-15 medium (Gibco), and further incubated in this medium. Cells were imaged using a Zeiss LSM 510 Meta inverted confocal microscope fitted with a heated stage unit for live-cell imaging (heating insert P, Zeiss). Heterochromatin or euchromatin regions, selected based on EGFP-HP1- β localization, were damaged by applying 200 iterations of 100% power from a 405-nm diode laser. To monitor changes, serial time-lapse images were taken at 256×256 pixel 8-bit resolution with a scan speed of 293 ms, using the 488-nm line from the argon laser at 1% power. For some experiments, cells were fixed at the indicated times for staining and immunofluorescence.

A pulsed N_2 laser system with a 365-nm dye cell (Micropoint; Photonics systems) coupled to the epifluorescence path of a Zeiss 510 LSM Confocor2 was also used for DNA damage experiments to induce localized damage in a large number of cells. The cycling time was set at approximately 10 Hz and the power used to excite Hoechst for DNA damage (calibrated by γ -H2AX staining)

was set at maximal attenuation of laser output through the supplied gradient neutral density filter (approximately 75%). Cells visualized under phase contrast were moved by manual operation of the motorized stage through a fixed point of laser illumination to induce stripes of DNA damage across several nuclei. Typically, 100–200 cells were damaged per experiment, and then fixed for immunofluorescence.

Immunofluorescence. Cell staining and immunofluorescence analyses were carried out as described previously². Briefly, transfected and untransfected cells were grown on coverslips and subjected to different treatments as indicated in the text before fixation. Cells were fixed using 4% formaldehyde for 10–15 min, followed by solubilization and blocking with 3% BSA in PBS, 0.05% Tween-20, 0.05% Triton. Imaging was performed on a Zeiss LSM 510 Meta confocal microscope, using a $\times 40$ objective with fixed optical slice, laser power and detector/amplifier settings for all samples across each individual experiment to allow comparison. Quantitative analysis for HP1- β (T51P) and γ -H2AX staining was performed using mean per-nucleus intensity to avoid any discrepancy in the definition or counting of foci. This was performed on identically stained and imaged samples by estimation of raw pixel intensity values for a defined number of cells, specified for each experiment, through the Zeiss LSM software, and comparison of the different groups by one-way ANOVA and Dunnett's post test. High-content microscopy was also used for certain experiments; 96-well plates (Nunc) with cells seeded at 3,000 per well were treated with specified DNA-damaging agents, pre-treatments/siRNA, then fixed and immunostained. The plates were imaged and analysed using a Cellomics high-content screening microscope, using a $\times 40$ non-immersion objective. Hoechst staining was used for object identification and the average intensity of fluorescent signal per nucleus was estimated using the target activation Cellomics bio-application. Five-hundred cells were analysed per well, and standard errors were calculated from an average of the means of eight wells.

Expression and purification of HP1- β . Full-length human HP1- β or its chromodomain were expressed in *Escherichia coli* BL21 (Stratagene) as carboxy-terminal fusions to GST using the pGEX4T3 vector (Pharmacia). Bacteria expressing the fusion protein were lysed in PBS supplemented with 5 mM dithiothreitol, 0.5% NP-40 and protease inhibitors, and the soluble fraction was applied onto glutathione-Sepharose 4B beads (Amersham). After extensive washing, full-length HP1- β and the chromodomain fragment were cleaved from the GST moiety using thrombin protease (Amersham) in PBS at 22 °C for 16 h. Thrombin was removed from the sample by chromatography with benzamidine-Sepharose 4B (Amersham).

Peptide synthesis and binding assay. A peptide (FITC-H3K9me2) encoding residues 1–15 of the H3 tail dimethylated on Lys 9 and amino-terminally labelled with fluorescein was synthesized by PeptideCrest Limited. Ten micrograms of purified GST-HP1 chromodomain (or its T51A mutant form), 3 nmol of FITC-H3K9me2 peptide (representing a tenfold molar excess over HP1) and 20 μl of 50% GST bead slurry were added to a PBS-based binding buffer containing 0.2% NP-40 and 0.1% BSA, and incubated at room temperature for 1 h. For phosphorylation experiments, 25 μl of a kinase reaction containing 10 μg GST-HP1 chromodomain protein, 5,000 units rCK2, in CK2 buffer containing 50 μM ATP, replaced the unmodified GST-HP1 chromodomain. After binding, the mixture was centrifuged at 5,000g for 3 min and all supernatant removed by gentle aspiration. The beads were re-suspended in 50 μl binding buffer, transferred to assigned wells of a 96-well plate, and fluorescein fluorescence intensity read using a Fusion microplate reader (PerkinElmer). A mixture of FITC-H3K9me2 and beads alone was used to calibrate background binding, to which the rest of the data were normalized before export to GraphPad Prism to calculate mean intensities and standard errors from different runs ($n = 3$). Changes in binding were shown to be significant using a student's t -test.

26. Daniels, M. J., Marson, A. & Venkitesan, A. R. PML bodies control the nuclear dynamics and function of the CHFR mitotic checkpoint protein. *Nature Struct. Mol. Biol.* 11, 1114–1121 (2004).

27. Rogakou, E. P., Boon, C., Redon, C. & Bonner, W. M. Megabase chromatin domains involved in DNA double-strand breaks *in vivo*. *J. Cell Biol.* 146, 905–916 (1999).

Exploring unseen communities

Advances in sequencing technology and tools for analysis are allowing researchers to unravel the environmental diversity of microbes faster and in greater detail than ever before. Nathan Blow reports.

This year marks the tenth birthday for metagenomics — the cloning and functional analysis of the collective genomes of previously unculturable soil microorganisms in an attempt to reconstruct and characterize individual community inhabitants. Since the term was coined by Jo Handelsman and her colleagues at the University of Wisconsin in Madison, its scope has expanded greatly with descriptions of the microbial inhabitants of environments as diverse as the human gut, the air over New York, the Sargasso Sea and honeybee colonies. And within these communities researchers are now uncovering a wider range of microorganisms, thanks in large part to advances in DNA-sequencing technology.

“We can look at the metagenomic analysis so much more deeply, at such a better cost,” says Jane Peterson, associate director of the Division of Extramural Research of the National Human Genome Research Institute in Bethesda, Maryland, which recently launched a five-year initiative to explore the human microbiome.

Although sequencing technology is creating opportunities for metagenomics research,



The 454 Life Sciences GS FLX sequencing system is used in many metagenomics projects.

all these new data are straining downstream analysis. “Computational analysis of metagenomic data still has quite a few outstanding questions,” says Isidore Rigoutsos, manager of the bioinformatics and pattern-discovery group at IBM’s Thomas J. Watson Research Center in Yorktown Heights, New York. The assembly and prediction of gene function for high-complexity microbial communities still poses challenges¹, for example (see ‘Benchmarks and standards’).

Maybe it is the promise of rapidly improving sequencing technology or the new environments being explored, but Peterson says that she has seen a growing interest in large metagenomics projects — particularly the Human Microbiome Project, which aims to unravel the microbial communities associated with various parts of the human body, including the gut (see page 578). “People somehow identify with the Human Microbiome Project. It is interesting how this project, especially as it is studying the gut, has really caught a lot of people’s attention.”

Over the past few years, the race to sequence DNA faster and more cheaply has been taking

454 LIFE SCIENCES

BENCHMARKS AND STANDARDS

The complexity of microbial communities can vary drastically, from a couple of microorganisms to thousands or even millions, making the reconstruction of whole genomes from some samples tricky. “If the community is low in complexity, it should allow one to reconstruct genomes with high accuracy,” says Isidore Rigoutsos, manager of the bioinformatics and pattern-discovery group at IBM’s Thomas J. Watson Research Center in Yorktown Heights, New York. But when it comes to highly complex communities, things are less straightforward.

Rigoutsos and his team have tested several genome assemblers and gene-prediction tools on simulated metagenomic data sets with varying degrees of complexity. Knowing the composition of the community allowed the team to benchmark and evaluate the tools.

“We found that as the complexity increased, many of

the computational tools had an increasingly hard time,” says Rigoutsos. For most high-complexity samples, he says, the genome assemblers could not generate larger contigs, and several contigs that were assembled were actually chimaeric mixtures of sequences.

For metagenomic analysis, smaller contigs and single reads make assigning the sequence to a specific microorganism difficult. “We want to be able to assign a read of less than 1,000 nucleotides,” says Rigoutsos, which might allow researchers to determine species composition from high-complexity samples without the need to generate larger contigs.

Rigoutsos and his colleagues have made three simulated data sets available to researchers interested in testing assembly and prediction programs.

The problem of data analysis is not restricted to metagenomics — a

growing number of researchers are using next-generation sequencing platforms and generating the quantity of data that in the past might only have been possible at large genome centres. Several companies are developing software to address this issue.

CLC bio in Cambridge, Massachusetts, offers the CLC Genomics Workbench, which provides reference assemblies of data from various next-generation sequencing systems as well as mutation detection. A future version of the program will incorporate algorithms for the *de novo* assembly of Sanger as well as next-generation sequence data. Meanwhile, Geospiza in Seattle, Washington, and GenomeQuest in Westborough, Massachusetts, are developing software to analyse data generated by Applied Biosystems SOLID next-generation sequencing platform.

The combination of assembly

software and data sets to benchmark results should help solve some of the complexity problems associated with metagenomics. “If you sequence sufficiently, even 200 base-pair reads are enough,” says Rigoutsos. But he adds that the real question is how many 200 base-pair reads will be needed before we can truly understand complex communities.

Others are finding that with enough reads, fewer than 200 base pairs might be sufficient. Jens Stoye from Bielefeld University in Germany has compared a data set of 35 base pair reads generated on the Genome Analyzer from Illumina in San Diego, California, with a 454 data set for the same low-complexity sample. Although 99% of the Genome Analyzer’s sequence data were discarded, because the system generates up to 50 million reads he could assign the species in the sample with the same efficiency from both data sets.

N.B.

centre stage. Several next-generation DNA sequencing systems are now available, boasting gigabase outputs for a variety of genetic applications. But when it comes to sequencing environmental samples that contain many different microorganisms in varying amounts, the next-generation options have their limitations.

"I would say the only next-generation sequencing technology suitable for metagenomics at the moment is the 454 system," says Stephan Schuster a biochemist at Pennsylvania State University in University Park. Schuster is not alone — almost all metagenomic studies currently being reported rely on either 454 technology or conventional Sanger sequencing. The main reason is simple: read length.

Long-term tool

Developed by 454 Life Sciences in Branford, Connecticut, the 454 system relies on an emulsion polymerase chain reaction (PCR) step that is coupled to pyrosequencing. Individual fragments of DNA, 300–500 base pairs long, are attached to beads *in vitro* and amplified with PCR to generate millions of identical copies on each bead. Fragments are then sequenced by use of a massively parallel reaction format in 1.6 million wells on a picotitre plate. Using this system, researchers can generate around 250 base pairs of sequence per reaction while performing 400,000 reads in a single instrument run.

Other next-generation sequencing systems typically generate fewer than 50 base pairs

per reaction, relying instead on more reads to generate a greater number of base calls. These shorter read lengths are suitable for applications such as candidate gene resequencing, transcriptional profiling and microRNA discovery. But the 454 system's longer read length has attracted metagenomics researchers, and is enticing the community away from Sanger sequencing, which provides reads of 700–1,000 base pairs per reaction.

Forest Rohwer, a microbiologist at San Diego State University in California, sees the advantages of 454 when analysing sequence data from environmental samples. "We know that when you are below 35 base pairs it is hard to get information out, at 100 base pairs you tend to lose information, but if you get to 200 base pairs you start to gain a lot of information," he says. Although he also adds that it is not clear at the moment how much could be gained from even longer reads.

"The information content might be a little different because the Sanger reads are longer. At present, 454 is between 200 and 400 base pairs, depending on who you talk to, so you are not likely to cover an open reading frame, but you do get a significant amount of data," says Karen Nelson, an investigator at the J. Craig Venter Institute in Rockville, Maryland. But



Karen Nelson is one of many investigators working on the Human Microbiome Project.

she notes that studies using 16S ribosomal RNA pose a problem for 454 read lengths.

16S ribosomal RNA is sequenced to give a sense of species abundance and composition in a community without reconstructing entire genomes from the inhabitants. It is around 1,500 base pairs long, which can be sequenced in two or three Sanger reads. But, according to Nelson, with the 454 system, researchers will have to focus on a variable region of the gene to identify the spe-

cies or come up with other metrics, instead of relying on full-length reads.

Nelson thinks one approach might be to combine Sanger capillary sequencing with 454, so that Sanger methodology could be used to sequence the ends of inserts from large-scale cosmid or fosmid libraries, which would act as scaffolds for the placement of the shorter, but more numerous, 454 reads.

But the length of 454 runs could soon be less of an issue. "We are getting very close to entire exons and whole genes with our next version," says Michael Egholm, vice-president of research and development at 454 Life Sciences, noting that the company plans to introduce a 500-base-pair capacity instrument later this year.

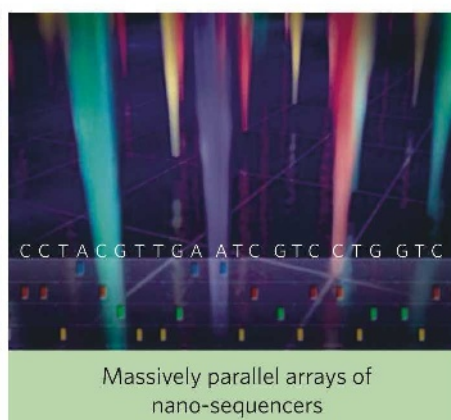
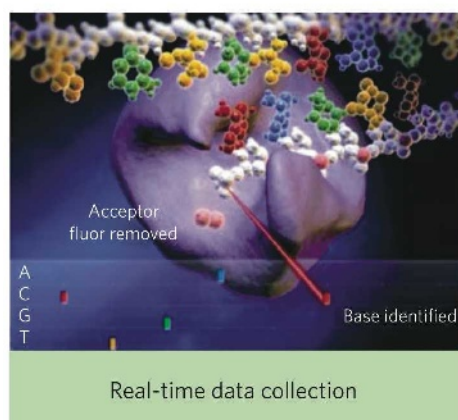
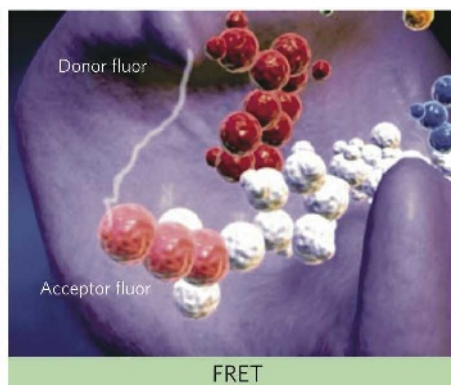
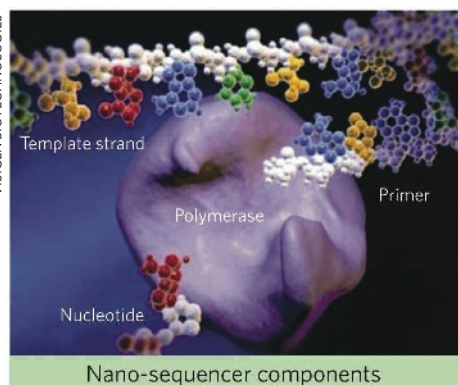
Diverse results

The widespread use of next-generation sequencing technology to explore microbial communities has brought something else to light — greater microorganism diversity. Next-generation systems bypass the cloning of DNA fragments before sequencing, a necessary step for most Sanger sequencing, and this has resulted in the discovery of new microorganisms that previously had been missed because of cloning difficulties. The lack of cloning has also made determining relative numbers of microbes in the community easier.

"As there is no cloning, we have very low bias," says Egholm. Schuster and others have also noted less bias when applying 454 sequencing to their environmental samples. "What is very comforting to see is that a lot of tests have been done by different groups finding the same results with 454 sequencing and qPCR," says Schuster.

Egholm and his colleagues are using their sequencing platform for several large-scale metagenomics projects — the biggest of which actually originated by chance. "The biggest metagenomic project on Earth might be our Neanderthal genome project," says Egholm. They are using 454 to sequence the complete genome of a Neanderthal, which Egholm says they hope to release by the end of the year. But 95–98% of the DNA in the Neanderthal sample comes from the environment rather than from

VISIGEN BIOTECHNOLOGIES



VisiGen Biotechnologies is developing a FRET-based approach to single-molecule sequencing.

a Neanderthal. This means that to get the 1× coverage, or roughly 3 billion base pairs, of the genome, the team must sequence somewhere between 70 billion to 100 billion base pairs of these environmental samples.

As Egholm's team begins to comb through those environmental data contaminating the Neanderthal samples, he says the initial results have been surprising. "Whether these are recent bacteria, or ones that ate the poor guy when he died, we cannot be certain yet, but I can tell you from the first few contigs where we got multi-kilobase lengths — they matched nothing in GenBank."

The next, next generation

"The biggest change in metagenomics will come from 'third generation' sequencing systems or single-molecule sequencing," says Schuster. For the metagenomics community this next-generation promises longer reads than Sanger sequencing, even higher throughput, lower costs and better quantitation of genes.

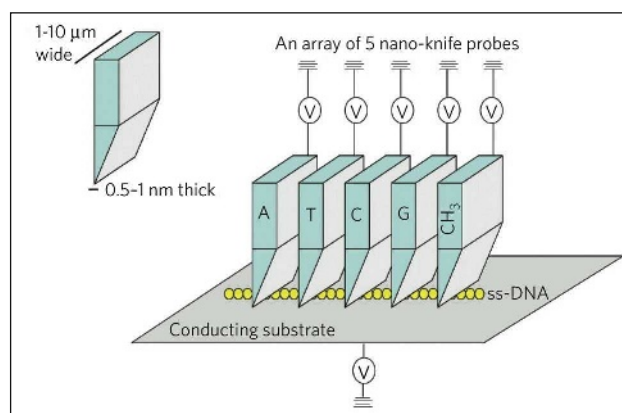
VisiGen Biotechnologies in Houston, Texas, is working on a method for sequencing single molecules in real-time that uses Förster resonance energy transfer (FRET). In this system, the polymerase is engineered to contain a donor fluorophore, and each nucleotide has a differently coloured acceptor fluorophore attached to the gamma

phosphate. When one of the nucleotides is incorporated by the polymerase into a native strand of DNA, a unique FRET signal is given off, and the pyrophosphate-containing fluorophore is then released leaving the synthesized strand of DNA ready for the next incorporation event. By imaging the colour changes as each incorporation occurs, VisiGen hopes to sequence single molecules in real-time and apply this on a massively parallel array with an output of up to one million bases per second.

Reveo in Hawthorne, New York, is developing a method to sequence DNA using nano-knife-edge probes, which pass over DNA that has been stretched and immobilized in a channel 10 micrometres wide. By using four different nano-knife-edge probes, each 'tuned' to a different frequency, Reveo hopes to non-destructively sequence DNA while eliminating the need for a costly imaging component. And Pacific Biosciences of Menlo Park, California, recently announced its single-molecule sequencing technology based on zero-mode waveguides².



Isidore Rigoutsos is testing genome assembly.



Reveo is developing a single-molecule sequencing technology that physically probes DNA.

These new single-molecule sequencing methods show the promise of generating much longer read lengths in the future, says Schuster. VisiGen's method predicts sequence reads of 1,500 base pairs — the size of 16S ribosomal RNA — whereas the Pacific Biosciences approach could produce reads as long as 10 kilobases.

And single-molecule sequencing is becoming a reality: in March 2008, Helicos BioSciences of Cambridge, Massachusetts, sold the first single-molecule sequencing system. The HeliScope uses a sequencing-by-synthesis approach in which the DNA is first fragmented into pieces 100–200 base pairs long. Adaptors of known sequence are attached to the ends of the fragments so that they can be captured on

THE HUMAN ENVIRONMENT

"I don't know why the Human Microbiome Project bubbled to the top now instead of previously," says Jane Peterson from the National Institutes of Health (NIH). "Certainly there have been metagenomic studies done with the old technologies." But Peterson thinks recent reports exploring the human gut microbial community, along with new, advanced sequencing technologies, might have been enough to pique the interest of the reviewers who decided to fund the Human Microbiome Project (HMP).

The project is a 5-year, US\$115-million effort to study the microbial communities inhabiting several regions of the human body, including the gastrointestinal and female urogenital tracts, oral cavity, nasal and pharyngeal tract, and skin, and how those communities influence human health and disease. The effort is

viewed by many researchers in the metagenomics community as particularly timely as most agree improvements in sample collection standards and analysis tools are much needed. Karen Nelson from the J. Craig Venter Institute, who is also a participating investigator in the HMP, says these issues can finally be addressed with a project of this scope as all samples will be collected and treated in the same way, and standards will be put in place for annotation of the metagenomic data.

The project will fund research in several areas, although the construction of a data resource for sequencing DNA samples from as many as 250 individuals, and projects aimed at demonstrating how changes in the human microbiome are related to health and disease are centrepieces of the programme. Other project initiatives include the development of new metagenomics technology

to isolate bacteria that are currently cannot be cultured, development of new bioinformatic programs and tools for analysis of large genomic data sets, data analysis and coordination centres, and analysing and understanding the ethical, legal and social issues of the project.

The HMP's initial sequencing efforts began this year. Peterson says that towards the end of the last fiscal year, the NIH Roadmap office had funds available for the HMP. Through existing relationships with large sequencing groups, the HMP was able to start quickly and generate preliminary sequencing data, which are being used for the other demonstration projects. This initial effort will result in the sequencing of 200 new bacterial organisms, recruitment of patients for metagenomic studies, and some 16S ribosomal RNA metagenomic sequencing to

assess microbial diversity at the various sites.

Peterson says that the protocols for sampling and recruitment have provided some early challenges. The HMP hopes to sample the same sites on all 250 individuals, but with so many sites, standardizing sampling can be tricky. "The protocol for the different sites has to be well worked out," she says, "the oral community has to be happy with the requirements that the skin community brings to the table."

Although it is just at the beginning, the HMP is scheduled to award its first rounds of grants to researchers this autumn. Peterson says that in the future the project's standardization efforts might not be restricted to the United States. "We are also forming an international consortium to coordinate international projects."

N.B.

the surface of a coated flow cell. Once attached, a mix of labelled nucleotides and polymerase is flowed through the cell and the surface is imaged at different times to determine whether labelled nucleotides have been incorporated. The key to the technology is a method to effectively cleave off the labelled nucleotides following incorporation, permitting additional rounds to be performed.

"The moment metagenomics goes down to the single-molecule level, it will be possible to assess even very low abundance messages at a very large sequence interval," says Schuster.

Information overload

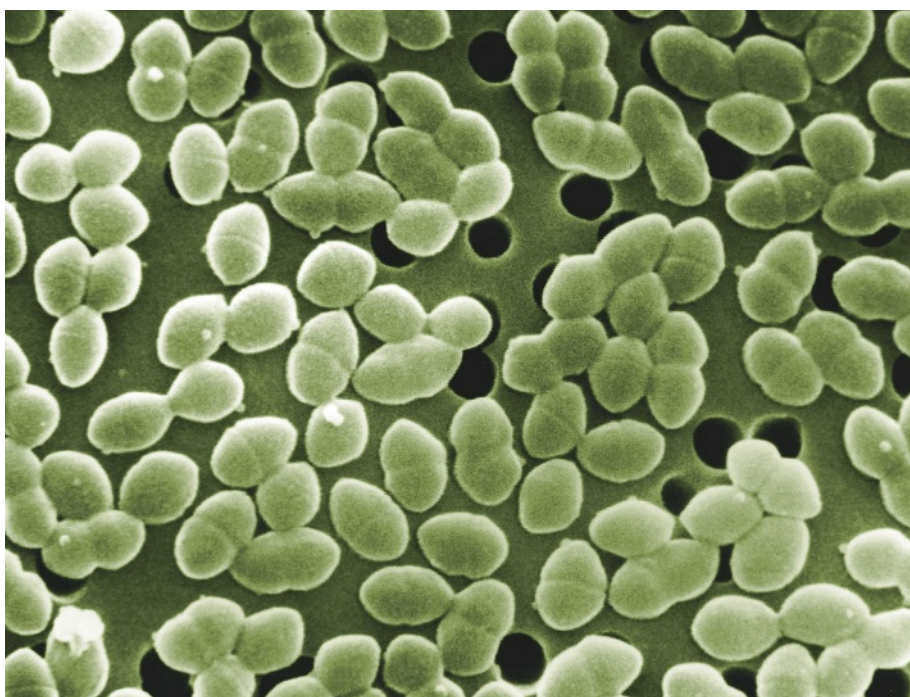
"People struggle at the moment with extracting basic information from metagenomic data," says Peer Bork a bioinformatician at the European Molecular Biology Laboratory in Heidelberg, Germany. Bork and others say that most bioinformatic analyses being done on metagenomic data sets involve relatively basic procedures such as assembly and gene annotation.

There are several options for such procedures. Some use web-based tools, such as the Rapid Annotations using Subsystems Technology (RAST) server developed by Argonne National Laboratory and the University of Chicago, in which a metagenomics data set is submitted and an analysis file is returned. Others, such as the Metagenome Analyzer (MEGAN) program, developed by Schuster and his colleagues, run on a desktop computer.

"I think being able to handle these large data sets and developing tools for visualization are critical, as they will allow researchers to write meaningful publications," says Schuster. He explains that MEGAN uses a binning format based on the National Center for Biotechnology Information's taxonomy database to display how often a certain taxon occurs. Other efforts, including the Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis, or CAMERA, aim to bring together bioinformatics resources and a data repository to assist with analysis of the data sets.

One issue with the various tools available, notes Rohwer, is that none has been adopted as standard. "It is still pretty much a cottage industry," he says.

"Standardization of annotations and gene-prediction quality during the early stages of metagenomic studies is something that needs to be addressed," agrees Nelson. She adds that it is now much easier to generate data than to interpret what they mean. "In a number of situations we are dealing with unknown species that have not had their genomes sequenced, so there will not be a reference genome to align to."



Enterococcus species are an example of bacteria found in the human gastrointestinal tract that will be examined during the Human Microbiome Project.

Rigoutsos agrees. "How can you tell the phylogenetic provenance of a sequence segment if the databases contain no examples of it?" he asks. He sees the situation as being akin to the early days of human genomics when some gene-prediction programs relied on database searches to draw conclusions. His group has developed PhlyoPythia, a software tool that also takes a binning approach to classifying sequence contigs assembled from metagenomic data sets.

Annotation is not the only standardization issue. "We need to do a better job of collecting information when we take samples," says Rohwer. The time, place and collection method can profoundly affect the microbial composition in a sample. Whether the sample is collected during the day or at night, acquired from a person's right or the left arm, or even if two soil samples are collected 5 millimetres apart, these

seemingly small differences can lead to very different communities of organisms. Rohwer thinks that collecting a standard set of information for each sample would make future comparisons between different data sets easier and so provide greater biological insight.

There is hope that some of these issues will be addressed as members of the metagenomics community become increasingly involved with large-scale, multi-institution projects (see 'The human environment'). "The Human Microbiome Project is going to be one good example

where standards are in place for data acquisition, generation and analysis," says Nelson.

Dynamic future

With bioinformatics tools and sequencing poised to go even faster at a lower cost, researchers are eyeing the next level of metagenomic analysis — a move from simply cataloguing the microbes in a community to understanding the interactions and dynamics of the organisms.

"I think a systems-level approach that gives a holistic view of these communities will open different research possibilities," says Rigoutsos, whose group is now working on studying metagenomes for biofuel applications.

Rohwer hopes to use metagenomics to manipulate a system and then trace how the community goes through changes at a global level. He thinks that with sequencing speed and cost declining so rapidly, these experiments, which only a few years ago would have been impossible, are now within reach.

Bork also wants to move towards analysing global communities. "Most groups concentrate on the early parts, but I think it is time to ask the bioinformatics people to develop methods beyond there, to get ecology concepts projected on those molecular data."

Nelson says that all the technology development and interest in metagenomics is spurring the field along nicely. "I think the technology is very promising and it will get better. It is a great time to be doing this."

Nathan Blow is the technology editor for Nature and Nature Methods.

1. Mavromatis, K. *et al.* *Nature Methods* **4**, 495–500 (2007).
2. Korlach, J. *et al.* *Proc. Natl Acad. Sci. USA* **105**, 1176–1181 (2008).



Forest Rohwer studies phage distribution in environmental samples.

| COMPANY | PRODUCTS/ACTIVITY | LOCATION | URL |
|---|---|-----------------------------|--|
| DNA sequencing and analysis | | | |
| 454 Life Sciences | Automated DNA analysis systems including Genome Sequencer 20 and FLX | Branford, Connecticut | www.454.com |
| Applied Biosystems | Automated sequencing systems using both Sanger methodology and next-generation SOLiD-based system | Foster City, California | www.appliedbiosystems.com ● |
| Beckman Coulter | Automated DNA sequencing systems based on capillary electrophoresis using the Sanger method | Fullerton, California | www.beckmancoulter.com |
| Bionanomatrix | Using nanofluidics platforms to develop new sequencing systems | Philadelphia, Pennsylvania | www.bionanomatrix.com |
| Complete Genomics | Developing sequencing technology based on a sequencing-by-hybridization approach using combinatorial probe-ligation chemistry | Sunnyvale, California | www.completegenomics.com |
| Cogenics | Contract automated DNA sequencing services using both Sanger-based and next-generation sequencing systems | Morrisville, North Carolina | www.cogenics.com |
| GE Healthcare | Kits for automated DNA sequencing systems | Buckinghamshire, UK | www.gehealthcare.com |
| Helicos BioSciences | Developing HeliScope Genetic Analyzer based on single-molecule sequencing methodology | Cambridge, Massachusetts | www.helicosbio.com |
| Illumina | Illumina Genome Analyzer next-next-generation DNA sequencing system | San Diego, California | www.illumina.com |
| Licor Biosciences | Automated DNA sequencing instrumentation using the Sanger methodology | Lincoln, Nebraska | www.licor.com |
| Pacific Biosciences | Developing automated next-generation sequencing instrument using zero-mode waveguides | Menlo Park, California | www.pacificbiosciences.com |
| Promega | Supplies the fMol DNA sequencing system for manual DNA sequencing and promoter mapping applications | Madison, Wisconsin | www.promega.com ● |
| Roche Applied Sciences | Supplies the 454 Life Sciences Genome Sequencer 20 and FLX systems | Basel, Switzerland | www.roche.com |
| Saturn Biotech | Full DNA sequencing service provider | Perth, Australia | www.saturnbiotech.com |
| VisiGen Biotechnologies | Developing next-next-generation DNA sequencing based on a real-time single-molecule approach | Houston, Texas | www.visigenbio.com |
| Zymo Research Corporation | DNA and RNA isolation kits; systems for sequencing reaction purification; kits for isolation of DNA from soil samples | Orange, California | www.zymoresearch.com |
| Bioinformatics tools for DNA sequence analysis | | | |
| CLC Bio | Developing algorithms and software for the analysis of next-generation sequencing data | Cambridge, Massachusetts | www.clcbio.com |
| DNASTar | Software for DNA sequence and microarray analysis | Madison, Wisconsin | www.dnastar.com |
| Geospiza | Software for management of genetic data | Seattle, Washington | www.geospiza.com |
| SeqWright | Contract automated DNA sequencing services | Houston, Texas | www.seqwright.com |
| Genetic and genomic testing companies | | | |
| 23andMe | Seeks to explore personal genomics | Mountain View, California | www.23andme.com |
| AncestryByDNA | Developing genetic testing products and services | Sarasota, Florida | www.ancestrybydna.com |
| Athena Diagnostics | Diagnostic testing services for neurological disorders | Worcester, Massachusetts | www.athenadiagnostics.com |
| DNADirect | Genetic testing services; genetic counselling | San Francisco, California | www.dnadirect.com |
| GeneDx | Genetic testing services; mutation confirmation services | Gaithersburg, Maryland | www.genedx.com |
| Genetic Technologies | Genetic testing services for people, animals and plants | Fitzroy, Australia | www.gtg.com |
| Genzyme Genetics | Human genetic testing services for reproductive genetics and oncology/pathology | Cambridge, Massachusetts | www.genzymegenetics.com |
| IdentiGEN | Genetic testing of products for the agriculture and food industries | Dublin, Ireland | www.identigen.com |
| Knome | Complete genome sequencing and analysis services for individuals | Cambridge, Massachusetts | www.knome.com |
| Gene expression analysis | | | |
| Agilent Technologies | Microarrays; lab-on-a-chip methods; SNP analysis software | Santa Clara, California | www.agilent.com |
| Genomic Solutions | Assay assembly for SNP genotyping; sequencing and PCR | Ann Arbor, Michigan | www.genomicsolutions.com |
| Operon Biotechnologies | Oligonucleotide production; preprinted oligonucleotide arrays and probe sets | Huntsville, Alabama | www.operon.com |
| Oxford Gene Technology | Oligonucleotide arrays for various applications | Oxford, UK | www.ogt.co.uk |
| Qiagen | DNA, RNA extraction kits; genotyping; haplotyping | Germantown, Maryland | www.qiagen.com |
| TeleChem International | Microarray robots, printers and scanners; microarray services | Sunnyvale, California | www.arrayit.com |
| Vermillion | Biomarker identification for drug discovery | Fremont, California | www.ciphergen.com |

| COMPANY | PRODUCTS/ACTIVITY | LOCATION | URL |
|-----------------------------------|--|-----------------------------|--|
| General | | | |
| Applied Biophysics | Automated instruments for cell monitoring and electric cell-substrate impedance sensing | Troy, New York | www.biophysics.com |
| Alexis Biochemicals | Suppliers of reagents for molecular- and cell-biology research | Lausanne, Switzerland | www.alexis-corp.com |
| Andor Technology | Microscopy, chemiluminescent kits, single-molecule detection, cameras and imaging solutions for FRET experiments, and total internal reflectance microscopy systems | Belfast, Ireland | www.andor.com |
| BioFlow Technology | Developing bioreactor systems for cell culture in a three-dimensional environment | Novi, Michigan | www.bioflowtech.com |
| Biomol | Services for chemical synthesis, cell culture and antibody production | Hamburg, Germany | www.biomol.de |
| Bio-Rad | Products, instruments and software for life-sciences research | Hercules, California | www.bio-rad.com |
| BMG Labtechnologies | Microplate and array readers and handling systems | Offenburg, Germany | www.bmglabtech.com |
| Brinkmann Instruments | Laboratory instrument supplier; consumables | Westbury, New York | www.brinkmann.com |
| Cambrex | Products for molecular- and cell-biology research | East Rutherford, New Jersey | www.cambrex.com |
| EMD Biosciences | Calbiochem, Novabiochem and Novagen product lines | San Diego, California | www.emdbiosciences.com |
| Enzo Life Sciences | Consumables and reagents for molecular biology, gene expression and genomic analysis | New York, New York | www.enzo.com |
| Eppendorf | Consumables for molecular biology; instrumentation | Hamburg, Germany | www.eppendorf.com |
| EUGENEX Biotechnologies | Development of test cell lines | Taegerwilen, Switzerland | www.eugenex.com |
| Geneservice | Genomic and proteomic resources including DNA samples, cDNA and genomic libraries; contract services including DNA sequencing, microarray analysis and SNP genotyping. | Cambridge, UK | www.geneservice.co.uk |
| Gilson | Pipettes, automated liquid handling, liquid chromatography systems and software | Middleton, Wisconsin | www.gilson.com |
| Hamilton Company | Automated liquid-handling stations | Reno, Nevada | www.hamiltoncompany.com |
| Harvard Apparatus | Instruments and equipment for electrophysiology and cell biology | Holliston, Massachusetts | www.harvardapparatus.com |
| Horiba Jobin Yvon | Spectroscopy systems and accessories including Raman, atomic emission and UV spectroscopy | Edison, New Jersey | www.jobinyvon.com |
| Integra Biosciences | Equipment for sterilization, liquid handling, cell culture and sample storage | Baar, Switzerland | www.integra-biosciences.com |
| Invitrogen | Kits and reagents for molecular-biology research; fluorescent probes for FRET experiments | Carlsbad, California | www.invitrogen.com |
| Irvine Scientific | Defined media for cell-culture applications; custom media services | Santa Ana, California | www.irvinesci.com |
| Merck | Chemicals, kits and reagents for molecular- and cell-biology-related research | Darmstadt, Germany | www.merck.de |
| Molecular Devices | Liquid-handling and microplate processing equipment; imaging instruments | Sunnyvale, California | www.moleculardevices.com |
| MP Biomedicals | Reagents and chemicals for research | Aurora, Ohio | www.mpbio.com |
| Nalge Nunc International | Labware | Rochester, New York | www.nalgenunc.com |
| New England Biolabs | Molecular-biology-related reagents, kits and enzymes | Ipswich, Massachusetts | www.neb.com |
| Opticell | Automated cell-culture devices | Westerville, Ohio | www.opticell.com |
| Perkin Elmer Life Sciences | Instruments, reagents and kits for life-sciences research | Waltham, Massachusetts | las.perkinelmer.com |
| Pierce Chemical | Protein assays, purification, Western blotting | Rockford, Illinois | www.piercenet.com |
| Premier Biosoft | Software for life-sciences research | Palo Alto, California | www.premierbiosoft.com |
| Roche Diagnostics | Reagents and kits for molecular biology; genomics instrumentations and software | Lewes, UK | www.roche-applied-science.com |
| Stratagene | Tools and reagents for molecular biology, genomics and proteomics | La Jolla, California | www.stratagene.com |
| Takara Bio | Reagents, kits and consumables for molecular biology | Shiga, Japan | www.takara-bio.com |
| Tocris Cookson | Chemicals for life-science research; contract research services | Avonmouth, UK | www.tocris.com |
| Wako USA | Speciality chemicals supplier; clinical diagnostic reagents | Richmond, Virginia | www.wakousa.com |
| USB | Chemicals and reagents for molecular biology | Cleveland, Ohio | www.usweb.com |

● see advertisement

naturejobs

**JOBS OF
THE WEEK**

What should be the role of the wise elderly scientist who, despite his or her age and failing pipetting skills, continues to make legitimate scientific contributions? Are senior-citizen scientists an asset, or are they obstacles to the hiring of new bright young minds?

In a Commentary in this week's issue (page 588), 66-year-old developmental biologist Peter Lawrence of the University of Cambridge, UK, staunchly defends the elderly scientist. He cites the long career of fly geneticist Seymour Benzer and the exploits of 80-year-old Nobel laureate Sydney Brenner, who was recently hired by the Howard Hughes Medical Institute's Janelia Farm campus in Ashburn, Virginia.

Lawrence criticizes the mandatory retirement ages that apply in many universities and for government employees in Europe. Not only is it discriminatory, he argues, but it impedes making commitments to graduate students and diminishes negotiating power. Some institutions do allow scientists to apply for extensions, but in Germany, for example, philanthropic foundations have had to start special initiatives to continue to fund eminent scientists due for retirement (see *Nature* **445**, 334–335; 2007). Meanwhile, the United States, with a law that bans mandatory retirement ages, has for years benefited from an 'elderly brain gain'. Accomplished scientists attracted across the Atlantic help to buoy many a university's standing.

But could less-capable elderly scientists hang around too long? Keeping on older scientists could, in principle, make funds and university slots harder to come by for younger scientists in a funding environment that is already difficult in places like the United States — although Lawrence dismisses this argument.

He does suggest some controversial remedies, however: scientists with retirement pensions should only get a salary supplement; tenure should not be indefinite; and each scientist should be regularly evaluated for his or her contribution and asked to go when they fail to measure up. But in the end, perhaps the only way is to let each individual decide: should I continue to trek to the lab, or should I be content to let the next generation take up the mantle?

Gene Russo is editor of Naturejobs.

CONTACTS

Editor: Gene Russo

European Head Office, London
The Macmillan Building,
4 Crinan Street, London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Andy Douglas (4975)
e-mail: a.douglas@nature.com
Business Development Manager:
Amelie Pequignot (4974)
e-mail: a.pequignot@nature.com
Natureevents:

Claudia Paulsen Young (+44 (0) 20 7014 4015)
e-mail: c.paulsenyoung@nature.com
France/Switzerland/Belgium:
Muriel Lestringuez (4994)
Southwest UK/RoW: Nils Moeller (4953)

Scandinavia/Spain/Portugal/Italy:
Evelina Rubio-Hakansson (4973)
Northeast UK/Ireland:
Matthew Ward (+44 (0) 20 7014 4059)
North Germany/The Netherlands:
Reya Silao (4970)
South Germany/Austria:
Hildi Rowland (+44 (0) 20 7014 4084)

Advertising Production Manager:
Stephen Russell
To send materials use London address above.
Tel: +44 (0) 20 7843 4816
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com
Naturejobs web development: Tom Hancock
Naturejobs online production: Dennis Chu

US Head Office, New York
75 Varick Street, 9th Floor,
New York, NY 10013-1917
Tel: +1 800 989 7718

Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

India
Vikas Chawla (+91 1242881057)
e-mail: v.chawla@nature.com

Japan Head Office, Tokyo
Chiyoda Building, 2-37 Ichigayatamachi,
Shinjuku-ku, Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:
Ayako Watanabe (+81 3 3267 8765)
e-mail: a.watanabe@natureasia.com
Business Development Manager, Greater China/Singapore:
Gloria To (+852 2811 7191)
e-mail: g.to@natureasia.com

MOVERS

Larry McKinney, director, Harte Research Institute for Gulf of Mexico Studies, Texas A&M University, Corpus Christi, Texas



1986–2008: Director of resource protection, senior director of aquatic resources, and director of coastal fisheries, Texas Parks and Wildlife Department, Austin, Texas

1980–86: Director, Texas Environmental Engineering Field Laboratory, Galveston, Texas

Larry McKinney grew up in a small town in west Texas during the worst drought in the state's history. It inspired him to revolutionize Texas's water policies by using ecosystem-based coastal management. Colleagues hope his new move will raise awareness of the Gulf of Mexico's national importance: more than 60% of US lands drain into the gulf.

Escaping the desert, McKinney became an oceanography undergraduate at Texas A&M University in Corpus Christi. His PhD focused on the origin and distribution of amphipods, tiny crustacea, throughout the Gulf and Caribbean. A summer fellowship at the Smithsonian Institution working with the biologists there sealed his academic intentions.

But first he found his expertise in demand for addressing real-world problems. Having returned to Texas A&M, McKinney became involved in several high-profile research projects. One — studying the environmental impacts of pumping concentrated brine out of oceanic salt zones to create cavities for the strategic petroleum reserve — led to the creation of the Texas Environmental Engineering Field Laboratory, which McKinney oversaw. He successfully brought in additional grants and contracts for other coast-based research work, but got burnt out chasing money.

When the Texas Parks and Wildlife Department asked him to help establish a resource-protection division, he accepted. Endangered-species concerns soon created controversy among landowners. McKinney crafted several policies, including funding for landowner conservation efforts, which became part of the federal Endangered Species Act.

At the Parks and Wildlife Department, McKinney sought to incorporate ecosystem-based management by connecting conservation, hydrology and economics. Andy Sansom, now director of the River Systems Institute at Texas State University in San Marcos, says McKinney successfully got an ecological component into Texas water policy despite the state legislature's reluctance. "Larry developed that capability and capacity by bringing sound science — that could not be discounted — to the table," says Sansom.

McKinney says two massive challenges face the Gulf of Mexico: population growth and climate change. "If we don't build a strong science base to develop policies while we still have some options, it's going to be a disaster," he says.

With McKinney ready to tackle such nationwide issues, Samson says the Harte Institute will realize its potential to become a Scripps Institution of Oceanography or Woods Hole Oceanographic Institute for 'America's third coast'. ■
Virginia Gewin

NETWORKS & SUPPORT

The softer side of science

Success in science is about more than mastering lab techniques. It also depends on 'soft' skills such as motivation, personality, research strategy and communication. It is not always easy for well-trained objective truth-seekers to consider soft skills, which are subjective. But they may help you boost your productivity and communicate your science better.

Scientists should shield themselves from discouraging events and develop a 'frustration tolerance' for paper and grant rejections, criticism by well-meaning colleagues and the depressing tedium of data collection. Then there's hypermotivation. To avoid burn-out, try relaxation (sports, yoga, meditation) and a healthy social life. Just as learning requires a quiet consolidation phase to store material in long-term memory, success in science needs intermittent silence.

Personality traits cannot usually be changed, but there are ways to improve one's disposition in the lab. Perfection, for example, can only be expected in pure mathematics or fairy tales. Beware of the 80–20 rule: for perfection, the last 20% of a task may take 80% of the effort — not a wise choice if you want to be productive. Worse, perfectionism is a sure path to leaving projects or papers incomplete.

A related trap is failure to bring a

project to publication. A finished experiment may satisfy your curiosity, but data are only of value to your CV and the rest of the world if published. The drive to completion is healthy, if you wish to succeed in science. Many of us have papers that are 90% finished but never submitted. Consider the time you have already invested, how little is left to do and how much effort it would take to get to the same point with another project.

Writer's block can be a major challenge: sitting in front of a blank page, lacking the wherewithal to start writing. This may be linked to the perfectionist trap ("I do not write unless my text is perfect"), but it could also be the result of a lack of ideas, of writing ability or of self-confidence. It is usually the first 10 minutes of writing that are hardest. Try to start writing without worrying about the presentation or structure.

Scientific leadership qualities — exhibiting responsibility, flexibility and trust — are also essential in all aspects of science.

These seemingly simple tasks can be taxing. But improving soft skills is a critical element of science success. ■

Bernhard Sabel is a member of the medical faculty at the Institute of Medical Psychology at the University of Magdeburg in Germany.

POSTDOC JOURNAL

Singapore's happiest person

My husband Andy is 'Singapore's happiest person'. He won a contest, organized by the aptly named Philip Merry, that was part of a conference on happiness and well-being. Judges selected the winner from nominations submitted by the public, based on criteria such as disposition, resilience and impact on others. But as the news headline puzzled: "What's he so jolly about?"

Many people in Singapore aspire to the 'five Cs': cash, credit card, car, condominium (privately developed luxury apartments) and country-club membership. Fierce drive and ambition pervade the working environment in Singapore, so much so that friends who relocate overseas do not know what to do with their free time; they fear that it made them less competitive.

Yet Andy's happiness defies this definition of success. We are humans, not automatons. Andy works hard and plays hard, focusing on four Fs: family, friends, fun and food. He likes organizing barbecues and dinner parties, at the end of which everyone rolls home satiated and happy.

A visiting researcher once emphasized the importance in science of openness and unselfishness. Outside the lab these qualities can contribute to a positive feedback cycle that buoys spirits. I try to emulate Andy's positive effect on others, in the lab and in general. I would not trade it for any amount of cash. ■

Amanda Goh is a postdoctoral fellow in cell biology under the Agency of Science, Technology and Research in Singapore.

Morpho sanguinalis

Beauty is in the eye of the beholder.

Julie Jansen

The first recorded sighting of *Morpho sanguinalis* in the Pacific Northwest occurred in 2020. Dr Fitzmorris identified it as he gazed through a pair of binoculars into his Seattle backyard. The world-renowned entomologist knew the South American butterfly well.

An unmistakable species, the *Morpho* was enormous, with brilliant azure wings that spanned 20 inches. Beauty was its only redeeming quality. Dr Fitzmorris had lost half his men to the insect during his last expedition to the Venezuelan jungle. The trip was a nightmare he would never forget. *Morpho sanguinalis* had a taste for human blood. With its mosquito-like proboscis, a group of *Morphos* could drain the blood from an adult human in less than an hour.

Dr Fitzmorris set the binoculars down. He rubbed his eyes and hoped it was only a hallucination. The stunning butterfly floated above a tree. On the ground below, blood dribbled from a cut on Mrs Fitzmorris's knee, injured while tending her garden. There soon appeared another *Morpho*, and then another, until 20 circled over his wife like hungry vultures.

"Run!" he screamed. She raised her head, her face shaded by the wide brim of a sunhat. *Morpho sanguinalis* had arrived, and Mrs Fitzmorris, along with the entire northwest population, was in trouble.

Climate change led to many new insect species calling the region their home: fire ants, killer bees, botflies. Although rainfall hadn't changed, within 20 years the area lost its winter chill and morphed into a hot, humid, tropical jungle. The climate was perfect for *Morpho sanguinalis*, but geographical barriers had kept them out. The butterfly couldn't have reached the area without human help.

The butterflies fluttered closer to his wife. He screamed at her again, but she only shrugged her shoulders, oblivious to the imminent danger. When her gaze drifted to the house her expression turned angry. Their daughter, Marny, stood outside the door in a risqué French maid's outfit clutching two large opened specimen boxes. Marny grinned as she watched the butterflies dance above her mother's head.

The horrible truth flashed before Fitzmorris's eyes. Marny was a troubled teen. She'd fallen in with a group of young women known as the Gothic Lolitas. Fitzmorris spent half his pay cheques on

outfits that transformed her into a black-lipstick-wearing Little Bo Peep. Perhaps not so coincidentally, the Gothic Lolitas worshipped the *Morpho*'s vampiric tendencies.

Marny had travelled with Dr Fitzmorris on the ill-fated trip to Venezuela. She wore those ridiculous outfits even in the jungle. Fitzmorris remembered her sewing the hem of her petticoats one night. He'd assumed she was mending a tear. Now he realized what she'd been doing — she was stitching *Morpho sanguinalis* pupae into her skirt, where they would go undetected by customs agents at the airport. Homeland security knew how to sniff out bombs and drugs, but pupal-stage insects slipped by scot-free.

Even in the adult stage, few would suspect the high level of danger the butterfly posed. More common was that people would gawk at its beauty until the *Morpho* swooped down and bit them on the neck. Toxins in the butterfly's saliva had a rapid paralysing effect. With the victim rendered immobile, one or more *Morphos* could gorge themselves with blood. Dr Fitzmorris and Marny witnessed it first hand in the Venezuelan jungle.

Marny and Mrs Fitzmorris argued often these days. The spats were getting more and more violent. They fought that morning because of the French maid's outfit Marny insisted on wearing to school. Mrs Fitzmorris forbade Marny to leave the house until she changed. Marny refused. Rather than obey her mother, the unruly teen unleashed her prized *Morphos*. They were attracted to the blood flowing from the cut on her mother's knee.

Mrs Fitzmorris noticed some dark shadows moving about the ground. She glanced up and saw what looked like several birds. The underside of *Morpho sanguinalis*' wings was a dull brown, a trait that kept it well-camouflaged against the trunks of trees. When one of the butterflies swooped in closer, Mrs Fitzmorris saw the brilliant blue wings.

Mrs Fitzmorris gasped. She gazed every day at a picture of a *Morpho*. Marny kept it as a screensaver. She remembered everything Dr Fitzmorris had told her about the vampiric South American butterfly. The

Morpho was repelled by garlic. Mrs Fitzmorris grew a patch of garlic in her garden. Pollinating insects loved it. Mrs Fitzmorris pulled a stalk from the ground, and squished the pungent bulb in her palm. She smeared the muddy garlic paste over her face and neck.

The effect was immediate. The *Morphos* smelt it and flitted away. Mrs Fitzmorris ran to the house and into the arms of her husband.

Marny, on the other hand, was infuriated. Her plan had backfired. Her mother was alive, and she had lost her



JACEY

prized pets. She ran into the garden with a butterfly net, but tripped over a sprinkler and scraped her knee. The butterflies smelled her blood before Marny even realized she was injured.

Twenty she had released in all. Dr Fitzmorris managed to capture eighteen. He kept them alive in his laboratory at the university, for research purposes. Marny, spared death thanks to her father and forgiven by both parents for her misdeed, visited the butterflies regularly. Dr Fitzmorris insisted she wear a lab coat over her ridiculous outfits.

Currently it is not known whether there is a breeding population of *Morpho sanguinalis* in the Pacific Northwest. An occasional sighting is reported, but no proof.

If you suspect you have seen or have been a victim of *Morpho sanguinalis*, and have lived to tell the tale, report to your local emergency room immediately. Mandatory quarantine and a series of blood transfusions will be necessary.

Julie Jansen is a resident of Olympia, Washington. Her stories have appeared in *The Harrow*, *The Horror Library* and *Necrology* magazines. She is co-editor of the 'zine *Multiple Lazy Voices* (www.multiplelazyvoices.com).